



Nouvelle Collection scientifique
Directeur : Émile Borel

Le Hasard

PAR

ÉMILE BOREL
Professeur à la Faculté des Sciences de Paris.

QUATRIÈME MILLE

LIBRAIRIE FÉLIX ALCAN

LE HASARD

DU MÊME AUTEUR

LIBRAIRIE FÉLIX ALCAN :

L'Aviation (*en collaboration avec* Paul Painlevé *et* Ch. Maurain), 8e édition, revue. 1 vol. in-16 de la *Nouvelle collection scientifique*. avec figures. . . . . 5 fr. 75

LIBRAIRIE GAUTHIER-VILLARS :

Leçons sur la théorie des fonctions (*Éléments de la théorie des ensembles et applications*), 2e édition revue et augmentée ; 1917 . . . . 7 fr. 50

Leçons sur les fonctions entières . . . . 3 fr. 50

Leçons sur les séries divergentes . . . . 4 fr. 50

Leçons sur les séries à termes positifs, *redigées par R. d'Adhémar*, 3 fr. 50

Leçons sur les fonctions méromorphes, *rédigées par L. Zoretti*. 3 fr. 50

Leçons sur les fonctions de variables réelles et les séries de polynomes, *rédigées par M. Fréchet*, avec des Notes par Paul Painlevé et Henri Lebesgue. . . . . 4 fr. 50

Leçons sur la théorie de la croissance, *rédigées par A. Denjoy*. 4 fr. 50

Introduction géométrique à quelques théories physiques, 1914. 5 fr. »

Leçons sur les fonctions monogènes uniformes d'une variable complexe, 1917. 7 fr. 50

LIBRAIRIE HERMANN ET FILS :

Éléments de la théorie des probabilités, 2e édit. ; 1910. 6 fr. »

LIBRAIRIE VUIBERT :

Introduction à la théorie des nombres et à l'algèbre supérieure (*en collaboration avec* Jules Drach). . . . . . 10 fr. »

LE HASARD

PAR

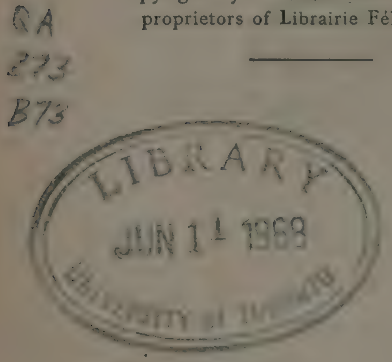
ÉMILE BOREL
Professeur à la Faculté des Sciences de Paris.

QUATRIÈME MILLE



PARIS
LIBRAIRIE FÉLIX ALCAN
108, BOULEVARD SAINT-GERMAIN, 108

1920


QA
273
B73
LIBRARY
JUN 1 1968

PRÉFACE

Il y a quelques mois, à une question sur mes travaux scientifiques, j'eus l'imprudence de répondre que je terminais un livre sur *le Hasard* ; mon interlocuteur me demanda aussitôt, non sans quelque ironie, ce que je dirais de neuf sur ce « magnifique sujet ».

Il me fallut assurer que je ne ferais connaître aucune recette pour gagner à la roulette, aucun talisman propre à écarter les hasards funestes ou à attirer sur nous et nos proches les hasards favorables ; mais, avec son mystère, le livre avait perdu tout son prestige et la question attendue : Mais que renferme-t-il donc alors ? fut posée avec une indifférence courtoise qui n'exigeait aucune réponse.

Les lecteurs habituels de cette Collection n'ont pas commis la même méprise et ne cherchent pas ici une sauvegarde contre les hasards de la vie. Ils pourront y trouver cependant, à côté de la spéculation pure, des remarques plus concrètes et même des

règles d'action pratique, pourvu qu'ils veuillent bien, pour les formuler, s'aider de leur réflexion personnelle.

La science du hasard, en effet, ne saurait, plus que tout autre science, prétendre à régir nos actes ; elle peut seulement, comme c'est le rôle de la science, faciliter la réflexion qui précède l'action chez tous les êtres raisonnables. Dans les questions compliquées, le bon sens a besoin d'être guidé par les résultats du calcul ; les formules ne créent pas l'esprit de finesse, mais en facilitent l'usage.

J'ai tâché de supprimer toutes les formules superflues ; celles qui subsistent contribueront, je l'espère, à préciser les idées des lecteurs qui ne sont pas entièrement brouillés avec l'algèbre ; elles peuvent être sans inconvénient omises par ceux qui sont habitués à croire les algébristes sur parole.

Mon but principal a été de mettre en évidence le rôle du hasard dans les branches diverses de la connaissance scientifique ; ce rôle a beaucoup grandi depuis un demi-siècle ; le moment est venu de nous demander si nous n'avons pas assisté, presque sans nous en apercevoir, à une véritable révolution scientifique.

La science a été dominée pendant deux siècles par la loi célèbre de Newton, qui a établi le règne de la mécanique rationnelle. C'est en prenant cette loi comme base ou comme modèle que les Lagrange, les Laplace, les Coulomb, les Gauss, les Fresnel, les Ampère, les Cauchy ont étendu le champ des explications mécaniques; mais, malgré tout leur génie et tous leurs efforts, le but suprême, l'explication mécanique de l'univers, s'est toujours dérobé, jusqu'au jour où la découverte et l'étude de la radioactivité ont montré que les explications mécaniques sont parfois certainement insuffisantes et doivent alors céder le pas aux explications statistiques.

Expliquer statistiquement un phénomène, c'est arriver à le regarder comme la résultante d'un très grand nombre de phénomènes inconnus régis par les lois du hasard. Si l'on arrivait à « expliquer » ainsi la loi de l'attraction universelle, on diminuerait le caractère mystérieux de cette loi, si belle par sa simplicité mais si absurde, il faut bien le dire, dans son énoncé classique d'après lequel l'attraction se transmet instantanément, sans intermédiaire, aux plus grandes distances. Ce problème est peut-être un des plus importants de la science actuelle.

Lorsque les applications aux sciences exactes auront perfectionné la théorie des probabilités, l'introduction des lois du hasard dans les sciences biologiques, sociologiques, psychologiques, etc., sera rendue à la fois plus aisée et plus féconde.

Par ces transformations de la science, la notion même de vérité scientifique se trouve modifiée. A l'affirmation dogmatique d'une loi comme celle de Newton est substituée la constatation de l'impossibilité pratique d'un hasard miraculeux. J'ai essayé, par l'exemple des singes pillant un dépôt de machines à écrire et reconstituant la Bibliothèque nationale en dactylographiant au petit bonheur, de rendre aussi concrète que possible cette notion du hasard miraculeux et cependant logiquement réalisable.

Je crois que ce mythe des singes dactylographes n'est pas sans avoir une réelle portée philosophique ; mais des développements étendus sur ce point auraient débordé le cadre de ce livre ; je me suis borné à quelques indications dont je prie qu'on excuse l'insuffisante brièveté.

Paris, février 1914.

PREMIÈRE PARTIE

LA DÉCOUVERTE DES LOIS DU HASARD

CHAPITRE PREMIER

LE HASARD ET LES LOIS NATURELLES

1-3. La nécessité des lois naturelles. — 4, 5. Le phénomène fortuit et la notion de probabilité. — 6. Critique de la définition de la probabilité. — 7. Probabilité objective et probabilité subjective. — 8. Le but de la théorie des probabilités.

1. — Une étude historique complète de la notion de loi naturelle se confondrait avec l'histoire de la raison humaine. La période la plus récente de cette histoire, celle de la genèse des sciences, est relativement bien connue, au moins dans ses grandes lignes ; nul n'ignore comment les spéculations abstraites des géomètres grecs ont été la source vive d'où sont sorties successivement la géométrie, l'algèbre, la mécanique, l'astronomie [1], la physique, et ont finalement rendu possible le développement industriel du XIXe siècle et l'asservissement chaque jour plus complet des forces naturelles au génie des héritiers de la pensée grecque. Le véritable ressort de cette splendide épopée, la conquête du globe par l'homme, c'est la foi en la raison humaine, la conviction que le monde n'est pas régi par les

1. Voir, au sujet de l'astronomie, l'excellent ouvrage de M. Jules Sageret : *Le Système du Monde, Des Chaldéens à Newton* (Paris F. Alcan).

dieux aveugles ou par le hasard, mais par des lois rationnelles : « ἀεὶ θεὸς γεωμετρεῖ » ; cette devise platonicienne signifie que le dieu qui gouverne l'Univers a une raison semblable à la raison des géomètres, c'est-à-dire que ceux-ci peuvent arriver à pénétrer les lois divines et immuables du monde : du jour où l'homme a compris qu'il pouvait se proposer un tel but, il ne s'en est jamais laissé détourner ; même dans les périodes les plus sombres de l'histoire, lorsque les soucis de la vie matérielle absorbaient presque toutes les énergies, il s'est trouvé des chevaliers servants de la raison pour conserver et transmettre le flambeau de la pensée antique.

2. — Il y a cependant toujours eu, il y a encore des métaphysiciens qui, habiles à dissimuler sous un langage obscur et prétentieux, ou bien élégant et imagé, le vide ou bien l'incohérence de leur pensée, ont prêché à l'homme le mépris de l'intelligence et de la raison. Nous chercherons, dans le dernier chapitre de ce livre, à préciser les cas où la raison humaine se trouve en face de problèmes trop compliqués pour elle et nous verrons quel peut être alors le rôle de l'intuition à côté de l'intelligence[1]; mais il n'est pas besoin de beaucoup de réflexion pour reconnaître que, sans leur intelligence, ceux qui se proclament anti-intellectualistes et exaltent l'énergie et la volonté, seraient vis-à-vis des « intellectuels » qu'ils méprisent dans la même situation qu'un indigène cou-

1. Sur le rôle nécessaire de l'intuition dans la découverte scientifique, je me permets de renvoyer à mes articles : La logique et l'intuition en mathématique (*Revue de métaphysique et de morale* 1907) ; La mécanique rationnelle et les physiciens (*Revue du Mois*, octobre 1910, t. X notamment, p. 421).

rageux s'attaquant à un européen timide, mais enfermé dans une casemate blindée et muni d'un excellent fusil[1] Ce qu'il ne faut pas oublier, c'est que l'intelligence doit se contenter de comprendre le monde et ne peut pas le créer : « *Mundum regunt numeri* » est un adage dont il ne faut pas abuser pour chercher à construire l'univers *a priori*, sans recourir à l'observation ; il faut l'interpréter : les nombres aident à comprendre le monde.

3. — Ce n'est pas seulement la science et ses applications industrielles qui nous imposent la notion de loi naturelle comme une condition nécessaire à notre existence ; l'impératif est encore plus catégorique, s'il est possible, lorsque l'on se place simplement sur le terrain de la vie pratique quotidienne. A ce point de vue, l'histoire de la période préscientifique de l'humanité serait encore plus instructive que l'histoire des deux ou trois derniers millénaires. Qu'il s'agisse de la culture du sol auquel on confie des semences précieuses en vue d'une récolte lointaine, ou des problèmes sans nombre qu'il a fallu résoudre pour l'élevage des animaux, la chasse, la pêche, la navigation, la conservation ou la cuisson des aliments, l'homme n'a pu vivre et progresser que grâce à la connaissance de lois naturelles toujours plus nombreuses et à une confiance grandissante en la valeur de ces lois. Quelque opinion métaphysique que l'on professe sur la contingence des lois

1. Je sais très bien que certains philosophes penseront que je simplifie vraiment trop la question ; je ne suis pas sûr qu'on ne puisse leur reprocher de l'avoir compliquée. Pour avoir le droit d'être anti-intellectuel, il faut d'abord renoncer à être intelligent ; sinon on commet un évident cercle vicieux.

de la nature, sur ce qu'elles étaient bien avant qu'il y eût des hommes et sur ce qu'elles seront longtemps après que l'humanité aura disparu (à supposer que ces questions aient un sens), il paraît incontestable qu'au point de vue pratique, pragmatiste comme on dit parfois, la croyance en ces lois est pour nous une nécessité : nous ne pourrions pas nous endormir si nous n'étions pas assurés que le soleil se lèvera demain. De même, on concevrait difficilement l'existence d'un homme qui, lâchant une pierre au-dessus de son pied, ne s'attendrait pas à la voir tomber et à avoir le pied écrasé.

La nécessité « humaine » des lois naturelles est le point de départ de toute spéculation scientifique ; ce principe est tellement évident qu'on juge inutile de le répéter dans chaque ouvrage de sciences ; j'aurais pu aussi le sous-entendre ; si je ne l'ai pas fait, c'est parce que le Hasard, qui est l'objet de ce livre, s'oppose précisément à la notion de loi ; il n'était donc peut-être pas inutile de rappeler brièvement la place éminente qu'il ne saurait être question de contester à cette notion.

4. — Malgré les progrès de la science, il y a beaucoup de faits que l'homme est incapable de prévoir ; l'exemple le plus banal est celui de la pluie et du beau temps ; nous n'y insisterons pas, car la complexité des phénomènes météorologiques ne nous permettrait d'en parler avec précision qu'au prix de trop longs développements. La sexualité dans les naissances humaines est un cas typique d'un problème jusqu'ici insoluble malgré la simplicité au moins apparente de son énoncé ;

tel enfant à naître sera-t-il un garcon ou une fille ? Nous ne savons même pas avec certitude si le sexe est déterminé dès l'instant de la conception ; la recherche des « causes » du sexe est un problème d'une complexité inouïe, bien qu'il n'y ait place que pour deux alternatives quant à la solution. D'autres phénomènes fortuits, tels que ceux qui se présentent dans les jeux de hasard (dés, cartes, etc.) sont d'une nature un peu moins complexe : qui analysera néanmoins en détail les mouvements de la main qui jette les dés ou qui bat les cartes ?

La caractéristique des phénomènes que nous appelons fortuits, ou dus au hasard, c'est de dépendre de causes trop complexes pour que nous puissions les connaître toutes et les étudier. Nous aurons l'occasion, dans le cours de cet ouvrage, de revenir à diverses reprises sur cette définition du hasard ; retenons-la simplement comme une première approximation, qui nous est suggérée par l'examen des conditions dans lesquelles se produisent les phénomènes les plus usuels.

Une étude, même très superficielle, des phénomènes fortuits les plus fréquents, montre que ces phénomènes obéissent à ce qu'on appelle des *lois statistiques*. Soient, par exemple, mille enfants à naître d'ici quelques mois ; pour chacun d'eux pris individuellement, nous n'avons actuellement aucun moyen de prévoir son sexe ; mais nous sommes sûrs de ne pas nous tromper en affirmant que sur les mille enfants, il y aura des garçons et il y aura aussi des filles. De même, si l'on donne un dé à chacun des soldats d'une armée et si chacun d'eux lance ce dé, on peut affirmer que le point 6 sera certainement amené par quelques-

uns. Sur un million de jeunes gens de vingt ans, reconnus d'excellente santé par les meilleurs médecins, on peut être certain que plusieurs seront morts avant dix ans ; et cependant chacun d'eux pris individuellement peut très légitimement espérer vivre jusqu'à trente ans.

On voit quelle est la différence qui sépare la loi statistique des lois naturelles : la loi statistique ne permet pas de prévoir un phénomène déterminé, mais énonce seulement un résultat global relatif à un assez grand nombre de phénomènes analogues ; de plus, sa certitude ne paraît pas être de même nature et n'entraîne pas l'assentiment de tous. Si, en présence d'une nombreuse assemblée, je prends une pierre dans la main et j'annonce que je vais la lâcher et qu'elle ne tombera pas, tous mes auditeurs resteront sceptiques et, si l'expérience annoncée réussit, chacun sera persuadé que la pierre a été maintenue en l'air par un fil invisible ou par tout autre moyen. Si, devant la même assemblée, je parie que je lancerai deux dés vingt fois de suite et que j'obtiendrai chaque fois le double-six, je rencontrerai certainement un égal scepticisme. Et, si l'expérience réussit, chacun croira à une supercherie. Les deux cas ne sont néanmoins pas identiques ; pour la pierre qui reste suspendue en l'air après que je l'ai lâchée, la certitude de la supercherie subsistera même après que j'aurai fait vérifier à chaque spectateur qu'on n'a pas employé les « trucs » auxquels il avait pensé et cette certitude sera assez fortement ancrée pour que chacun soit prêt à parier tout ce que l'on voudra, même sa propre vie, que l'expérience réussie n'est pas celle que j'avais décrite. Au contraire,

pour les vingt coups de dés, si les dés sont examinés avec soin, si l'on a confiance en mon honnêteté et si j'affirme que je les ai lancés loyalement, certains se demanderont peut-être si la coïncidence des vingt coups identiques n'est pas simplement due à un hasard sans doute fort invraisemblable, mais tout de même possible.

La question de savoir si cette impression est justifiée est au nombre de celles que nous chercherons à éclaircir plus loin ; mais le fait qu'elle existe ne peut pas être entièrement négligé ; il prouve que la loi statistique ne s'impose pas à l'esprit humain avec le même caractère de nécessité que les lois naturelles.

5. — Lorsque l'on étudie avec soin les phénomènes qui obéissent à des lois statistiques, on constate qu'il est possible de définir entre ces phénomènes des relations numériques assez régulières, la régularité apparaissant d'autant mieux que le nombre des phénomènes est plus considérable. Par exemple, si pendant plusieurs années l'on étudie la sexualité dans les naissances d'une petite agglomération humaine, on constate qu'il naît à peu près autant de filles que de garçons. Cette première observation conduit à penser que la naissance d'un garçon et la naissance d'une fille sont *également probables ;* on dira aussi que chacun de ces événements a pour lui *une chance sur deux*, ou que sa probabilité est $\frac{1}{2}$. Une observation plus prolongée ou étendue à une population plus nombreuse montre d'ailleurs qu'il naît en moyenne un peu plus de garçons que de filles ; en moyenne, sur 100 nais-

sances, il y aura 51 garçons et 49 filles ; on dira que la probabilité de la naissance d'un garçon est 0,51 et la probabilité de la naissance d'une fille 0,49.

La probabilité que nous venons de définir est ce qu'on appelle la probabilité statistique, parce que sa valeur n'est pas connue à l'avance d'après la nature du phénomène, mais résulte de la connaissance détaillée et précise d'un grand nombre de phénomènes. Certains auteurs insistent beaucoup sur la différence entre ces probabilités statistiques, qui sont de simples rapports entre des nombres observés, et les probabilités qu'il convient de définir d'une manière plus précise pour les soumettre au calcul. Comme exemple de ces dernières, lançons un dé bien fabriqué ; chacune des faces est également probable ; comme il y en a 6, la possibilité d'amener l'une d'elles est un sur six, c'est-à-dire $\frac{1}{6}$. Nous verrons que la différence entre les deux cas est plus apparente que réelle ; en réalité, toute probabilité concrète est, en définitive, une probabilité statistique définie seulement avec une certaine approximation. Bien entendu, il est loisible aux mathématiciens, pour la commodité de leurs raisonnements et de leurs calculs, d'introduire des probabilités rigoureusement égales à des nombres simples, bien définis ; c'est la condition même de l'application des mathématiques à toute question concrète : on remplace les données réelles toujours inexactement connues par des valeurs approchées sur lesquelles on calcule comme si elles étaient exactes : le résultat est approché, de même que les données [1].

1. Voir plus loin, p. 224.

En réalité, pour savoir si un dé est bien fabriqué, il n'y a pas d'autre moyen que de le lancer un grand nombre de fois et de constater que chacune des 6 faces revient à peu près aussi souvent que les autres[1].

En définitive, si parmi un grand nombre d'événements analogues, nous en distinguons certains sous le nom de *cas favorables*, nous dirons que *la probabilité est le rapport du nombre des cas favorables au nombre total des événements*. Si le cas favorable est la naissance d'un garçon et s'il est né 51.200 garçons sur 100.000 naissances, la probabilité de la naissance d'un garçon est 0,512.

6. — La définition que nous venons de donner est classique, mais sa critique n'est pas moins classique : on a souvent affirmé qu'elle renferme un cercle vicieux. Elle exige, en effet, la connaissance préalable de tous les événements possibles ; elle ne saurait donc servir à rien en ce qui concerne les événements futurs.

Ce cercle vicieux est au fond le même qui a été souvent reproché au syllogisme ; la majeure ne peut être affirmée que si tous les cas ont été examinés ; la conclusion était donc connue. Si j'affirme que *Tous les hommes sont mortels* et si j'en conclus que *Paul*, étant un homme, *est mortel*, mon affirmation initiale n'était légitime que si je savais déjà que Paul est mortel. Il n'y a pas lieu de discuter à nouveau ici le fondement

1. On pourrait, bien entendu, fabriquer par des moyens mécaniques très précis un dé identique à un autre qu'on aurait préalablement reconnu bien fabriqué ; la difficulté ne serait que reculée et compliquée.

de l'induction et la notion de loi naturelle ; nous avons rappelé que l'homme ne peut vivre que s'il croit à un grand nombre de majeures de syllogismes dont il n'a pas expérimenté tous les cas.

De même, nous ne savons pas combien il y aura de garçons sur les 100.000 premières naissances françaises de l'année prochaine ; mais, si nous avons constaté, par plusieurs observations successives, que sur 100.000 naissances françaises, il y a environ 51.000 garçons, nous sommes portés à croire qu'il en sera de même l'an prochain. L'étude de la théorie des probabilités permet de préciser la valeur de cette induction ; mais cette étude même repose sur un certain nombre d'inductions analogues, de même que les vérités scientifiques ont pour origine première des observations vulgaires et des inductions grossières.

7. — On a aussi critiqué la définition de la probabilité au moyen de la distinction entre probabilité objective et probabilité subjective [1]. Sans discuter pour le moment la légitimité de cette distinction, retenons simplement cette observation souvent faite : le hasard n'est que le nom donné à notre ignorance ; pour un être omniscient, la probabilité n'existerait pas. On peut observer de même que pour un être omniscient, toute la science et toute l'activité humaines seraient vaines et sans but ; ce n'est pas pour cet être que les hommes ont créé la science et la théorie des probabilités ; c'est pour eux-mêmes, qui sont loin d'être omniscients. Quels que soient

1. Voir plus loin, n° 87.

les progrès des connaissances humaines, il y aura toujours place pour l'ignorance et par suite pour le hasard et la probabilité.

Il faut toutefois se mettre en garde contre une confusion qui a été souvent une source de paradoxes. Si le hasard n'est autre chose que le nom donné à l'ignorance, la théorie des probabilités a donc pour but de créer de la science avec de l'ignorance, c'est-à-dire quelque chose avec rien. Il y a là une confusion sur le sens du mot ignorance. Voici un jeu de cartes qui a été longuement battu par plusieurs personnes ; j'y choisis une carte sans la regarder ; j'ignore évidemment complètement quelle est cette carte. Mais je sais néanmoins bien des choses concernant la petite expérience qui a consisté à extraire cette carte du jeu. Je connais tout d'abord la composition du jeu, soit que je l'aie vérifiée avant le battage des cartes, soit qu'il s'agisse d'un jeu neuf garanti par la marque d'une maison sérieuse ; si j'avais affaire à un prestidigitateur, je pourrais craindre que le jeu dont il use ne soit composé que de rois de carreau ; je sais ou je crois savoir que ce n'est pas le cas. Je sais aussi que les diverses personnes qui ont battu les cartes ont procédé loyalement, de leur mieux ; je les ai peut-être battues moi-même et je sais que j'ai procédé en tâchant de varier le plus possible mes procédés de battage, de sorte que je ne crains pas d'avoir laissé subsister l'ordre particulier qui pouvait exister avant le battage. Je sais aussi qu'en choisissant une carte sans la retourner, je n'ai été guidé par aucun signe m'en faisant connaître la nature. C'est sur tout ce que *je sais*, et non sur mon ignorance, que je baserai l'affirmation que la

probabilité pour que la carte choisie soit le sept de pique est précisément $\frac{1}{32}$, le jeu étant de 32 cartes. Si un comparse indiscret m'avertit que la carte choisie n'est pas une figure, comme mon ignorance sera diminuée et que l'hésitation ne portera plus que sur les 20 cartes qui ne sont pas des figures, la probabilité pour que ce soit le sept de pique sera $\frac{1}{20}$. Si l'on me dit que la carte est un sept, la probabilité sera $\frac{1}{4}$; et si enfin je retourne la carte, il n'y aura plus de probabilité, mais certitude positive ou négative. L'ignorance est donc sans aucun doute un élément nécessaire à l'existence de la probabilité, mais ce n'est pas l'ignorance seule qui crée la probabilité.

On peut observer à ce sujet qu'il peut y avoir plusieurs échelons dans l'ignorance : la superposition de plusieurs ignorances successives complique l'analyse des problèmes, sans introduire rien de réellement nouveau dans les principes ; aussi n'y a-t-il pas lieu d'insister ici ; donnons-en cependant un exemple simple, à titre d'indication, en laissant au lecteur le soin de l'étudier si cela l'intéresse : on a battu devant Paul dix jeux de cartes dont 4 sont des jeux de 52 cartes et 6 des jeux de 32 cartes ; on lui fait choisir une carte dans un de ces dix jeux sans qu'il puisse savoir s'il s'agit d'un jeu de 32 ou d'un jeu de 52 ; quelle est la probabilité pour que la carte choisie soit un roi de pique[1] ?

1. On observera que la probabilité n'est pas la même que si l'on choisissait au hasard une carte dans le jeu obtenu en battant ensemble les dix jeux ; elle est plus grande dans le cas du texte.

8. — Nous sommes maintenant à même de comprendre quels peuvent être l'objet et le but de la théorie des probabilités. On dit souvent que ce but est la recherche et l'étude des *lois du hasard ;* mais le hasard peut-il avoir des lois ? C'est l'objection que nous avons déjà signalée et que ce langage suggère invinciblement. Il est préférable de dire que *la théorie des probabilités a pour objet d'évaluer les probabilités d'événements complexes au moyen des probabilités supposées connues d'autres événements plus simples.* Son but, c'est d'arriver à prévoir avec une certitude presque absolue, humainement absolue peut-on dire, certains événements dont la probabilité est telle qu'elle se confond avec la certitude. Dans quelle mesure ce but peut être atteint, c'est ce que montrera toute la suite de cet ouvrage.

Si on le supposait pleinement réalisé, on voit que la théorie des probabilités compléterait singulièrement la connaissance que l'homme peut avoir de l'univers, puisque là où nous ne connaissons pas de lois naturelles et où règne pour nous le hasard, la connaissance des lois du hasard permettrait de remplacer les lois naturelles inconnues par d'autres lois aussi rigoureuses et aussi bien connues. Hâtons-nous de dire que cet idéal, non seulement n'est pas atteint, mais ne le sera pas tant que la science humaine ne sera pas achevée, c'est-à-dire jamais ; mais c'est déjà beaucoup que la théorie des probabilités permette d'étendre notablement notre connaissance de l'univers et notre capacité de prévision des phénomènes naturels. Nous allons essayer de faire comprendre, sur un exemple aussi simple que

possible, comment on peut, en combinant des probabilités simples qui laissent une large place à l'ignorance et au hasard, aboutir dans certains cas à des énoncés où la part de l'ignorance et du hasard se trouve avoir presque complètement disparu.



CHAPITRE II

LES LOIS DU JEU DE PILE OU FACE

9. Définition du jeu. — 10. Etude de 2, 3, 4 parties. — 11. Cas général : triangle arithmétique. — 12. Application. — 13. Cas des grands nombres. — 14, 15. Unité décimale d'écart ; son emploi. — 16. Loi des grands nombres de Bernoulli. — 17. Objections et paradoxes. — 18. La confusion entre la probabilité très faible et la probabilité nulle. — 19. L'objection de M. Le Dantec. — 20. L'allure d'une longue partie de pile ou face. — 21. Le gain ne peut pas être proportionnel au temps. — 22. La portée pratique de la loi des grands nombres.

9. — On jette en l'air une pièce de monnaie et l'on engage des paris sur le côté qui sera apparent après sa chute : l'un des côtés s'appelle pile et l'autre face. Tel est le problème le plus simple de probabilités, si l'on ajoute l'hypothèse que les chances sont égales pour le côté pile et le côté face. Au sujet de cette égalité des chances, nous laisserons de côté ici toute discussion philosophique, la prenant simplement pour un fait expérimental ou, si l'on veut, pour la *définition* même du fait que la pièce employée est *bonne* pour ce jeu : nous admettons dès lors que l'on se sert d'une pièce bonne. Si quelque lecteur pense que le relief irrégulier de la pièce ne permettra jamais qu'elle soit rigoureusement *bonne*, il peut se figurer que l'on emploie un jeton dont la symétrie a été rigoureusement recherchée par une fabrication des plus soigneuses. Cette symétrie

BOREL. — 2

parfaite[1] ne peut d'ailleurs être qu'un cas limite, une notion abstraite comme la notion de la ligne droite. Mais cette hypothèse peut être suffisamment réalisée dans la pratique pour que les conclusions auxquelles elle conduit s'accordent avec l'expérience, au degré d'approximation même des erreurs expérimentales : c'est tout ce que l'on peut demander à une science quelconque.

10. — Etudions d'abord les cas où l'on joue un petit nombre de parties et cherchons à voir quelles sont les circonstances qui pourront se produire. La première partie peut donner comme résultat pile ou face ; ces deux hypothèses sont également probables ; on convient de dire que la *probabilité* de chacune d'elles est *un* sur *deux*, ce que nous écrirons 1 : 2 ou 0,5. Nous pouvons représenter ce résultat par le tableau suivant

$$\begin{cases} P & 1 : 2 \\ F & 1 : 2 \end{cases}$$

Quel qu'ait été le résultat de la première partie, la seconde partie peut donner comme résultat pile ou face, et chacune de ces hypothèses est également probable, car le résultat de la première partie est sans influence sur celui de la seconde : suivant l'heureuse expression de Joseph Bertrand, la pièce de monnaie n'a ni conscience ni mémoire. J'insiste un peu sur ce point, malgré son évidence, car si on ne le comprend point d'une manière parfaite, sans restriction ni réti-

1. A parler en toute rigueur, la symétrie absolue entraînerait l'indiscernabilité des deux côtés, et par suite l'impossibilité du jeu ; si les deux côtés diffèrent en quoi que ce soit, ne serait-ce que par la couleur, on ne peut jamais être rigoureusement sûr que cette différence est sans influence.

cence, il est inutile de poursuivre l'étude de la théorie des probabilités.

C'est surtout l'habitude des jeux de hasard qui rend certains esprits réfractaires à cette notion de l'indépendance des événements successifs; comme ils ont observé que, dans une longue série, les coups de pile ou face sont à peu près également nombreux [1], ils en concluent qu'une longue série de coups ayant amené pile *doit* être suivie d'un coup face; c'est une *dette* que le jeu a contractée envers eux. Il suffit d'un peu de réflexion pour se convaincre à quel point cet anthropomorphisme est puéril : les raisons pour lesquelles les chances de pile et de face sont égales subsistent à chaque partie et l'on ne peut concevoir aucun mécanisme par lequel les résultats des parties antérieures pourraient modifier l'égalité des chances. Cette croyance anthropomorphique à la mémoire et à la conscience de la pièce de monnaie n'a donc aucun fondement positif; pour les esprits superstitieux, cela pourrait ne pas suffire à la condamner s'ils n'apercevaient pas d'autre moyen de rendre compte des observations qui donnent des nombres sensiblement égaux pour pile et pour face, sur un grand nombre de parties; mais nous allons précisément voir que l'on arrive à expliquer de la manière la plus satisfaisante ces résultats de l'expérience, en admettant l'indépendance des parties successives : tout prétexte disparaît donc pour contester cette indépendance.

1. L'observation des joueurs porte sur le rouge ou le noir à la roulette, ou sur le gain ou la perte du banquier au baccara, etc.; nous laissons de côté ces jeux, parce qu'ils sont plus complexes que le jeu de pile ou face.

Revenons donc à notre seconde partie; en combinant les résultats qu'elle peut donner avec les résultats de la première, nous obtenons le tableau suivant :

P	P
	F
F	P
	F

dans lequel la première colonne correspond à la première partie et la seconde colonne à la seconde partie; nous voyons que l'ensemble des deux parties conduit à quatre combinaisons possibles, toutes quatre également probables; on dira que la probabilité de chacune d'elles est *un quart*, c'est-à-dire 1 : 4 ou 0,25. On peut d'ailleurs les écrire comme il suit

PP	1 : 4	
PF	1 : 4	2 : 4 = 1 : 2
FP	1 : 4	
FF	1 : 4	

Nous avons réuni par un trait les deux combinaisons PF et FP ; elles ont en effet ceci de commun qu'elles renferment chacune une fois pile et une fois face ; *elles sont identiques quand on ne tient pas compte de l'ordre ;* si l'on énonce le résultat brut des deux parties, on peut dire que trois hypothèses seulement sont possibles : ou les deux parties amènent pile, ou une seule amène pile, ou aucune n'amène pile ; en d'autres termes, le joueur qui aurait joué pile peut gagner, ou 2 fois, ou 1 fois, ou 0 fois. Les probabilités respectives de ces hypothèses ne sont pas égales, comme une absence totale de réflexion aurait pu le laisser croire ; le cas où le joueur gagne 1 fois

sur 2 parties peut se produire par les deux combinaisons différentes PF et FP ; sa probabilité est donc *un demi*, tandis que la probabilité de chacun des autres cas est seulement *un quart*.

Supposons maintenant que l'on joue une troisième partie ; un raisonnement analogue nous conduira au tableau suivant

P	P	P
		F
	F	P
		F
F	P	P
		F
	F	P
		F

dans lequel la première colonne correspond à la première partie, la seconde colonne à la seconde partie, la troisième colonne à la troisième partie ; on voit qu'il y a en tout 8 combinaisons dont chacune est aussi probable que les autres. Modifiant l'ordre du tableau précédent, nous intervertirons la quatrième et la cinquième lignes, pour rapprocher les combinaisons qui conduisent au même résultat global, abstraction faite de l'ordre des parties, et nous obtiendrons ainsi le tableau suivant :

PPP	1 : 8	
PPF	1 : 8	
PFP	1 : 8	3 : 8
FPP	1 : 8	
PFF	1 : 8	
FPF	1 : 8	3 : 8
FFP	1 : 8	
FFF	1 : 8	

On voit que le joueur qui joue pile, par exemple, a une chance sur 8 de gagner 3 parties ; 3 chances sur 8 d'en gagner 2 ; 3 chances sur 8 d'en gagner 1, et une chance sur 8 d'en gagner 0.

On formerait de la même manière les tableaux relatifs au cas de quatre parties ; nous donnerons seulement le second des tableaux, où, pour la commodité de la lecture, nous remplaçons la lettre F par un point.

4P	PPPP	1 : 16	
3P	PPP.	1 : 16	4 : 16 = 1 : 4
	PP.P	1 : 16	
	P.PP	1 : 16	
	.PPP	1 : 16	
2P	PP..	1 : 16	6 : 16 = 3 : 8
	P.P.	1 : 16	
	P..P	1 : 16	
	.PP.	1 : 16	
	.P.P	1 : 16	
	..PP	1 : 16	
1P	P...	1 : 16	4 : 16 = 1 : 4
	.P..	1 : 16	
	..P.	1 : 16	
	...P	1 : 16	
0P	....	1 : 16	

On voit que le joueur qui joue pile a 1 chance sur 16 de gagner les 4 parties ; 4 chances sur 16 (ou 1 sur 4) d'en gagner 3 ; 6 chances sur 16 (ou 3 sur 8) d'en gagner 2 ; 4 chances sur 16 (ou 1 sur 4) d'en gagner 1 ; 1 chance sur 16 de n'en gagner aucune.

Sans qu'il soit nécessaire de pousser plus loin cette étude, on voit par quel mécanisme simple on arrive, en partant de probabilités supposées égales pour chaque partie, à trouver des probabilités inégales pour les divers

résultats globaux possibles d'un ensemble de parties. Par exemple, dans le cas de 4 parties les combinaisons dans lesquelles le résultat global est 2 fois pile et 2 fois face sont au nombre de 6; chacune d'elles, prise individuellement, n'est pas plus probable que la combinaison qui amène pile 4 fois de suite ; mais, lorsque l'on considère leur ensemble, il y a 6 fois plus de chances pour obtenir l'une d'elles, non désignée à l'avance, que pour obtenir la combinaison unique PPPP. Si deux joueurs font fréquemment des séries de 4 parties à pile ou face et s'ils observent avec soin les résultats du jeu, ils constateront sans peine que les cas les plus fréquents, au point de vue du résultat global, sont ceux où chacun d'eux gagne deux parties ; ces cas devront se produire en moyenne 6 fois sur 16; les cas où l'un d'eux gagne 3 parties se produiront 4 fois sur 16 pour chacun d'eux (et par suite 8 fois sur 16 si l'on ne spécifie pas lequel des deux gagne les trois parties; ils sont alors plus nombreux que les cas où l'égalité se produit); enfin il arrivera seulement une fois sur 16 que l'un d'eux gagnera les 4 parties (et par suite 2 fois sur 16 que les 4 parties seront gagnées par le même joueur, non désigné d'avance). Mais ce serait une très grave erreur de conclure de ces constatations que la combinaison PPPF, par exemple, est plus probable que la combinaison PPPP, c'est-à-dire que, les trois premières parties ayant donné pile, la quatrième a une tendance particulière à donner face. Si les combinaisons donnant pile 3 fois sont plus probables que la combinaison donnant pile 4 fois, ce n'est pas parce que chacune d'elles est plus probable, mais uniquement parce qu'elles sont plus nom-

breuses : il y en a 4, car la partie unique donnant face peut être soit la première, soit la seconde, soit la troisième, soit la quatrième ; c'est seulement ce dernier cas qui peut se produire lorsque les trois premières parties ont donné pile.

11. — Lorsque l'on considère un grand nombre de parties successives de pile ou face, il est aisé de donner une règle faisant connaître la probabilité d'amener m fois pile sur n parties. Il est clair que le nombre total des combinaisons possibles est 2^n, et que chacune de ces combinaisons est également probable car à chacune des combinaisons possibles pour les $n - 1$ premières parties en correspondent 2 pour les n parties, puisque la $n^{\text{ième}}$ partie peut donner soit pile, soit face ; c'est ainsi que nous avons trouvé 2 combinaisons dans le cas d'une partie, 4 dans le cas de deux parties, 8 dans le cas de trois parties, etc.

Parmi ces 2^n combinaisons, combien y en a-t-il renfermant m fois pile ? Nous désignerons leur nombre par C_n^m et nous remarquerons que, pour obtenir m fois pile, il faut ou bien que les $n - 1$ premières épreuves aient donné m fois pile et que la $n^{\text{ième}}$ donne face, ou bien que les $n - 1$ premières épreuves aient donné $m - 1$ fois pile et que la $n^{\text{ième}}$ donne pile. La première alternative fournit autant de combinaisons qu'il en existe donnant m fois pile sur $n - 1$ épreuves, c'est-à-dire C_{n-1}^m combinaisons ; de même, la seconde alternative fournit C_{n-1}^{m-1} combinaisons, on a donc

$$C_n^m = C_{n-1}^{m-1} + C_{n-1}^m.$$

On déduit de cette formule la règle suivante,

qui conduit au triangle arithmétique de Pascal[1].

On inscrit sur une première ligne le nombre 1 répété deux fois ; puis on calcule chaque nombre à inscrire dans les lignes suivantes en ajoutant au nombre inscrit immédiatement au-dessus de lui le nombre inscrit à la gauche de celui auquel on l'ajoute ; dans l'application de cette règle, on suppose mentalement que les lignes sont prolongées à droite et à gauche par des zéros. On obtient ainsi le tableau suivant :

1	1									
1	2	1								
1	3	3	1							
1	4	6	4	1						
1	5	10	10	5	1					
1	6	15	20	15	6	1				
1	7	21	35	35	21	7	1			
1	8	28	56	70	56	28	8	1		
1	9	36	84	126	126	84	36	9	1	
1	10	45	120	210	252	210	120	45	10	1

qu'il serait possible de prolonger et qui est de la plus grande utilité pour la solution des problèmes simples de probabilités. Les 4 premières lignes renferment les nombres que nous avons déjà trouvés au n° 10 dans les cas de 1, 2, 3, 4 parties ; la $5^{\text{ième}}$ ligne s'obtient par le calcul suivant

$$\begin{aligned} 0 + 1 &= 1 \\ 1 + 4 &= 5 \\ 4 + 6 &= 10 \\ 6 + 4 &= 10 \\ 4 + 1 &= 5 \\ 1 + 0 &= 1. \end{aligned}$$

1. Le *Traité du triangle arithmétique* est reproduit (texte latin et traduction française) dans le tome III des œuvres de Blaise Pascal, publiées par Léon Brunschvicg et Pierre Boutroux (Collection *les grands écrivains de la France*).

La somme des nombres de cette cinquième ligne est, d'après ces égalités mêmes, double de la somme des nombres de la quatrième ; elle est effectivement égale à 32 = 2 × 16. Il y a donc 32 combinaisons possibles pour 5 parties, parmi lesquelles 1 donnant 5 fois pile, 5 donnant 4 fois pile, 10 donnant 3 fois pile, 10 donnant 2 fois pile, 5 donnant 1 fois pile, 1 donnant 0 fois pile.

On peut remarquer que les diverses lignes du tableau sont symétriques, c'est-à-dire que les nombres équidistants des extrêmes sont égaux. Dans les lignes de rang pair, correspondant à un nombre pair de parties, les nombres vont en croissant jusqu'au milieu ; le nombre le plus élevé correspond à un partage égal des résultats entre pile et face ; dans les lignes de rang impair ce partage égal ne peut être réalisé et il y a au milieu de la ligne deux nombres égaux, plus grands que les autres et correspondant aux deux cas où l'un des deux côtés de la pièce apparaît une fois de plus que l'autre.

12. — Pour donner un exemple de l'emploi du triangle arithmétique, résolvons un problème élémentaire.

Problème. — *Un joueur à pile ou face a-t-il plus de chances de gagner 3 parties sur 4 ou d'en gagner 5 sur 8 ?* Il convient de préciser cet énoncé en indiquant si le joueur doit gagner précisément 3 parties et 5 parties ou s'il doit gagner *au moins* 3 parties ou 5 parties ; en d'autres termes, si le joueur gagne 6 parties, doit-on considérer qu'il en a gagné 5 ? C'est évidemment là une question de convention ; nous examinerons successivement les deux cas.

Premier cas : l'énoncé est pris au sens strict, c'est-à-dire que les nombres de parties gagnées doivent être précisément 3 et 5. On voit immédiatement sur le triangle arithmétique que, pour gagner 3 parties sur 4 il y a 4 cas favorables sur 16 et que pour gagner 5 parties sur 8, il y a 56 cas favorables sur 256 ; dans la première hypothèse, la probabilité de gagner est donc égale à $\frac{4}{16}$, c'est-à-dire à $\frac{1}{4}$, dans la seconde hypothèse, elle est $\frac{56}{256} = \frac{7}{32}$; elle est donc moins élevée ; *on a plus d'avantage à parier que l'on gagnera* 3 *parties sur* 4 *qu'à parier que l'on gagnera* 5 *parties sur* 8.

Second cas : l'énoncé est pris au sens large, c'est-à-dire que les nombres de parties gagnées doivent être au moins égaux à 3 et à 5. Le triangle arithmétique montre que pour gagner 4 ou 3 parties sur 4, il y a un nombre de cas favorables égal à $1 + 4 = 5$ sur 16 ; pour gagner 8, 7, 6 ou 5 parties sur 8, il y a un nombre de cas favorables égal à $1 + 8 + 28 + 56 = 93$ sur 256 ; dans la première hypothèse, la probabilité de gagner est $\frac{5}{16}$ ou $\frac{80}{256}$; dans la seconde hypothèse, elle est $\frac{93}{256}$, c'est-à-dire plus élevée ; *on a plus d'avantage à parier que l'on gagnera au moins* 5 *parties sur* 8 *qu'à parier qu'on en gagnera au moins* 3 *sur* 4.

13. — Le triangle arithmétique suffit pour résoudre les problèmes qui portent sur un petit nombre de parties ; mais il n'en est plus de même lorsque le nombre des parties devient considé-

rable ; pour 100 parties, il serait déjà malaisé de prolonger le triangle arithmétique jusqu'à la centième ligne ; pour 1.000 parties, ce serait à peu près impossible ; pour un million de parties, la tâche dépasserait les forces humaines.

L'algèbre et l'analyse mathématique fournissent le moyen d'obtenir assez aisément le résultat de calculs arithmétiques dont l'exécution directe serait trop longue ou même pratiquement impossible. Les formules que l'on obtient en simplifiant ainsi les calculs auxquels conduirait la théorie du triangle arithmétique, sont de la plus haute importance pour toutes les applications de la théorie des probabilités ; leur démonstration exige des développements mathématiques pour lesquels nous renverrons aux traités spéciaux sur les probabilités [1]. Le seul point sur lequel il importe d'insister ici, c'est que cette démonstration ne fait appel à aucun principe nouveau, ni à aucun postulat caché ; c'est simplement un mode de calcul abrégé, qui conduit exactement aux mêmes résultats que l'on obtiendrait si l'on avait le temps et la patience de prolonger suffisamment le triangle arithmétique ; ou, si l'on préfère, de recommencer un nombre suffisant de fois les raisonnements élémentaires du n° 10, en écrivant dans chaque cas le tableau complet des coups possibles.

14. — Pour rendre la formule des grands nombres aussi claire que possible, prenons un exemple précis. Soit une partie de pile ou face dans laquelle on a joué un million de coups. Il est aisé

1. Voir, par exemple, Emile Borel, *Eléments de la théorie des probabilités* (Hermann).

de se rendre compte que les combinaisons dans lesquelles il y a eu *exactement* le même nombre de fois pile ou face (c'est-à-dire 500.000 fois) sont relativement peu nombreuses par rapport au nombre total des combinaisons. Leur nombre diffère en effet très peu du nombre de celles qui comportent 500.001 fois pile et 499.999 fois face. Les formules qui font connaître les probabilités de chacune de ces combinaisons individuelles, peuvent se déduire des formules que nous allons donner; mais nous ne nous y attarderons pas, car elles sont très peu usitées; le problème que l'on se pose le plus généralement est le suivant : quelle est la probabilité pour que le nombre de fois où l'on amènera pile (ou face) dépasse de plus de 1.000, ou de plus de 20.000 le nombre 500.000 que l'on aurait amené s'il y avait eu partage égal ? En d'autres termes, quelle est la probabilité pour que le nombre des coups amenant pile soit compris entre 480.000 et 520.000 ? On dira, dans ce dernier cas, que l'*écart* est inférieur à 20.000, l'*écart* étant, par définition, la différence entre le nombre observé et le nombre que l'on obtient en multipliant le nombre total des coups (1.000.000) par la probabilité de pile $\left(\frac{1}{2}\right)$. Pour simplifier la formule, on modifie encore légèrement cet énoncé et on se propose de déterminer *la probabilité pour que l'écart dépasse un nombre donné.* Cette probabilité est évidemment égale au quotient du nombre des combinaisons dans lesquelles l'écart dépasse ce nombre donné, par le nombre total des combinaisons possibles; c'est ainsi qu'on le calculerait dans les cas simples, au moyen du triangle arithmétique. Le résultat que nous

allons donner est identique à celui qui se déduirait du triangle arithmétique dans le cas des grands nombres, si l'on pouvait exécuter le calcul complet.

Nous définirons d'abord ce que nous appellerons l'*unité décimale* d'écart; c'est l'écart tel que la probabilité pour qu'il soit dépassé soit précisément égale à *un dixième* [1]; dans le cas d'un grand nombre de parties de pile ou face, cette unité décimale d'écart peut être prise égale à la racine carrée du nombre des parties [2]; dans

1. Dans mes *Eléments de la théorie des probabilités*, j'ai introduit systématiquement l'*unité d'écart* que l'on peut appeler *népérienne* parce qu'elle est avec le nombre e dans la même relation que l'unité décimale avec le nombre 10.

2. J'omets ici, pour simplifier, certains détails qui n'ont d'intérêt que pour les calculs de précision, calculs très rarement nécessaires dans les applications usuelles et dont nous n'aurons pas besoin dans cet ouvrage. Pour de tels calculs, on recourra aux tables de la fonction habituellement désignée par $\Theta(\lambda)$ et définie par la formule

$$\Theta(\lambda) = \frac{2}{\sqrt{\pi}} \int_0^\lambda e^{-\lambda^2} d\lambda.$$

La probabilité pour que l'écart relatif λ soit dépassé est

$$1 - \Theta(\lambda) = \frac{2}{\sqrt{\pi}} \int_\lambda^\infty e^{-\lambda^2} d\lambda.$$

Les approximations que nous faisons consistent : 1° à remplacer par l'unité le facteur $\frac{2}{\sqrt{\pi}}$, qui est égal à 1,12 ; 2° à remplacer l'intégrale de $e^{-\lambda^2}$ depuis λ jusqu'à l'infini par $e^{-\lambda^2}$: cette seconde approximation est de signe contraire à la précédente ; 3° en remplaçant $e^{-\lambda^2}$ par $10^{-\mu^2}$, ce qui donne entre λ et μ la relation $\frac{\lambda}{\mu} = \sqrt{\frac{\log 10}{\log e}}$, à prendre pour valeur de ce rapport $\sqrt{2}$, ce qui revient à confondre $e = 2{,}78$ avec $\sqrt{10} = 3{,}1$. Dans le cas de m parties de pile ou face, l'unité d'écart népérienne est $\sqrt{\frac{m}{2}}$; en la multipliant par $\sqrt{2}$, suivant notre approximation 3°, nous obtenons bien $\sqrt{m}$.

le cas d'un million de parties, elle est donc égale à 1.000. La règle très suffisamment approchée pour la plupart des applications pratiques est alors la suivante : *La probabilité pour que l'écart soit égal à* n *fois l'unité décimale d'écart est* 10^{-n^2}. Par exemple, sur un million de parties de pile ou face, la probabilité pour amener plus de 501.000 fois pile (ou face) est 0,1, la probabilité d'un écart double, c'est-à-dire de plus de 502.000 pile (ou face) est 10^{-4}, c'est-à-dire un dix-millième ; la probabilité d'un écart triple (ou plus de 503.000) est 10^{-9}, c'est-à-dire un milliardième ; la probabilité d'un écart décuple, c'est-à-dire plus de 510.000 fois pile ou face est 10^{-100}, c'est-à-dire un nombre décimal dans lequel la virgule serait suivie de 99 zéros après lesquels serait écrit le chiffre 1. On voit avec quelle rapidité fantastique décroissent les probabilités dès que l'écart dépasse la valeur que nous avons appelée l'*unité décimale*[1].

En présence de ces résultats qui surprennent l'imagination, il n'est pas inutile d'insister à nouveau sur le fait que les méthodes qui y conduisent sont seulement *des méthodes de calcul abrégées*, mais que les mêmes résultats seraient

1. Pour donner, une fois pour toutes, une idée de l'approximation des formules simplifiées que nous donnons, voici les résultats exacts auxquels conduit l'emploi des tables de la fonction $\Theta(\lambda)$. L'écart dont la probabilité est $\frac{1}{10}$ n'est pas 1 000, mais 822 ; l'écart dont la probabilité est un dix-millième n'est pas exactement 2 000, mais 1 945 : l'écart dont la probabilité est un milliardième n'est pas exactement 3 000, mais environ 3 020. Pour les valeurs plus grandes de l'écart, les valeurs obtenues pour la probabilité sont si faibles qu'elles ne se prêtent à aucune application pratique ; et, d'autre part, au point de vue de notre imagination, une probabilité 10^{-95} ne se distingue pas d'une probabilité 10^{-100}.

obtenus par l'emploi du triangle arithmétique et des raisonnements élémentaires qui y conduisent, si l'on avait suffisamment de patience et de temps. C'est de la même manière que l'emploi des logarithmes simplifie les calculs d'intérêts composés pour de longues périodes, mais que les mêmes résultats seraient obtenus par le calcul élémentaire des intérêts simples, et leur addition au capital, année par année. L'analogie entre les deux problèmes est d'ailleurs très grande; dans le premier, on obtient des nombres considérables par la multiplication un grand nombre de fois par lui-même d'un facteur très peu supérieur à l'unité; dans le second, on obtient des nombres extrêmement petits par la multiplication de facteurs très peu inférieurs à l'unité[1].

15. — Les calculs analytiques par lesquels on arrive à la règle que nous avons donnée, supposent le nombre des parties assez grand, disons par exemple supérieur à 100; pour des petits

1. Dans le cas des intérêts composés, ces facteurs peu différents de l'unité sont tous égaux entre eux; dans le cas des probabilités, ils sont variables, et c'est cette variabilité qui introduit comme exposant le *carré* du rapport à l'unité décimale d'écart. Dans le cas d'un million de parties de pile ou face, la probabilité pour amener 500 001 fois pile est à la probabilité pour l'amener 500 000 fois dans le rapport $\frac{500\,000}{500\,001}$; pour passer à la probabilité pour amener 500 002 fois pile, il faut multiplier par $\frac{499\,999}{500\,002}$; pour passer à la probabilité pour amener 500 003 fois, il faut multiplier encore par $\frac{499\,998}{500\,003}$, et ainsi de suite; on conçoit qu'après un millier de multiplications on obtienne un résultat assez petit, résultat qui décroîtra d'autant plus vite dès lors que les facteurs tels que $\frac{499\,001}{501\,000}$ différeront déjà notablement de l'unité.

nombres, on arrivera à un résultat plus précis en calculant effectivement le nombre des combinaisons par le triangle arithmétique. Nous donnerons plus loin (n° 106) un exemple d'un tel calcul pour le cas de 20 parties. Il n'est pas inutile d'observer que, même dans le cas des petits nombres, si l'on désire connaître seulement l'ordre de grandeur du résultat, la règle de l'unité décimale conduit à une approximation suffisante, et a l'avantage d'être d'une application très aisée.

Par exemple dans le cas de 20 parties, l'unité décimale $\sqrt{20}$ est comprise entre 4 et 5 ; les nombres donnés au n° 106 montrent que la probabilité d'un écart égal ou supérieur à 4 est

$$\frac{2 \times 60.460}{1.048.576},$$

c'est-à-dire à peu près $\frac{3}{25}$, tandis que la probabilité d'un écart égal ou supérieur à 5 est

$$\frac{2 \times 21.700}{1.048.576},$$

c'est-à-dire environ $\frac{1}{25}$; la probabilité $\frac{1}{10}$ est bien comprise entre ces deux nombres. Cet exemple montre quel est le degré d'exactitude que l'on peut attendre de la règle de l'unité décimale d'écart. Cette exactitude est tout à fait suffisante lorsque, comme ce sera toujours le cas dans ce livre, on ne recherche pas les applications numériques précises qui sont indispensables, par exemple, à une compagnie d'assurances, mais l'on se propose simplement de se rendre compte de l'ordre de grandeur des probabilités qui interviennent dans les questions étudiées.

En raison de son importance, je rappelle la règle de l'unité décimale, pour le cas simple du jeu de pile ou face ; nous compléterons plus loin cette règle de manière à la rendre applicable à des problèmes plus généraux.

RÈGLE : *L'unité décimale d'écart est égale à la racine carrée du nombre des parties ; la probabilité pour que l'écart*[1] *dépasse* n *fois cette unité est* 10^{-n^2}.

16. — De cette règle on déduit aisément la démonstration de la proposition à laquelle Jacques Bernoulli a donné le nom de *loi des grands nombres.*

Pour un million de parties l'unité décimale d'écart est 1.000, pour cent millions de parties, elle est 10.000. La probabilité d'un écart égale à sept fois l'unité, par exemple, est dans les deux cas 10^{-49}. Dans le premier cas, cet écart correspond à plus de 507.000 parties gagnées ou perdues par le joueur qui parie pile, dans le second cas cet écart correspond à 50.070.000 parties gagnées ou perdues ; on voit que, lorsque le nombre des parties augmente, l'*écart absolu* que l'on peut craindre ou espérer augmente aussi, mais ce que l'on peut appeler l'*écart relatif*, c'est-à-dire le rapport de l'écart absolu au nombre des parties[2],

1. Pour parler rigoureusement, il faudrait dire *la valeur absolue de l'écart*. Mais la distinction est de peu d'importance pour les applications que nous ferons.

2. Désignons par a^2 le nombre des parties ; l'unité décimale d'écart est alors a ; la probabilité d'un écart na est donc e^{-n^2} ; mais quel que soit le nombre fixe n, si a est assez grand, na est très petit par rapport à a^2 ; donc, la probabilité d'un écart *relatif* aussi petit que l'on veut tend vers zéro lorsque le nombre des épreuves augmente indéfiniment.

va en diminuant. Sur un million de parties la probabilité pour que le rapport du nombre des parties gagnées ou perdues au nombre total des parties soit supérieur au quotient de 507.000 par un million, c'est-à-dire à 0,507 est égale à 10^{-49}, c'est-à-dire est extrêmement faible et pratiquement négligeable; c'est une éventualité qui ne se produira pas en moyenne une fois pendant des milliards de milliards de siècles si chaque habitant de la terre faisait des milliards d'expériences par seconde[1]. On trouvera la même probabilité pour que, sur cent millions de parties, le rapport entre le nombre des parties gagnées ou perdues et le nombre total soit supérieur au quotient de 50.070.000 par cent millions, c'est-à-dire à 0,5007. Si le nombre des parties était 10 milliards, l'unité décimale d'écart serait 100.000 et le rapport dont la probabilité serait toujours la même 0,50007. On voit que l'on peut parier avec d'au-

1. On est souvent amené à faire des comparaisons de ce genre pour rendre sensible à l'imagination l'extrême petitesse des nombres tels que 10^{-49}; indiquons brièvement les calculs qui justifient de telles comparaisons. Le nombre des habitants de la terre est inférieur à 10 milliards, c'est-à-dire à 10^{10}; il y a dans la journée moins de 100 000 secondes, c'est-à-dire moins de 10^5; en un siècle, il y a moins de 10^5 jours. Si donc, pendant un milliard de milliards de siècles, soit 10^{18} siècles, chaque habitant du globe faisait dix milliards d'expériences par seconde, soit 10^{10}, le nombre total de ces expériences serait inférieur à

$$10^{10} \times 10^5 \times 10^5 \times 10^{18} \times 10^{10} = 10^{48}.$$

Dans le cas où la probabilité pour qu'une expérience particulière conduise à un certain résultat est 10^{-49}, la probabilité pour que ce résultat soit obtenu une fois par l'ensemble de ces innombrables expériences sera à peu près $\frac{1}{10}$ (rigoureusement, c'est l'espérance mathématique que l'on obtient ainsi, car il faudrait tenir compte des cas où l'expérience réussit plus d'une fois; l'erreur est ici assez faible pour être négligée).

tant plus de certitude que ce rapport diffère très peu de 0,5, que le nombre des parties est plus grand. C'est la loi des grands nombres. Elle n'exprime pas autre chose qu'une propriété des combinaisons : lorsque le nombre des lettres P et F devient très grand, le nombre des combinaisons dans lesquelles ces lettres figurent à peu près aussi souvent l'une que l'autre croît prodigieusement plus vite que le nombre des combinaisons dans lesquelles elles figurent de manière inégale.

Nous verrons au chapitre IV, consacré aux probabilités des causes, comment la loi de Bernoulli, obtenue à partir de la notion de probabilité élémentaire, peut servir à son tour à éclaircir et à préciser cette notion.

17. — J'ai déjà indiqué[1] pour quelles raisons il ne me paraît pas inutile de faire accepter par le plus possible d'esprits cultivés certaines conclusions de la théorie des probabilités, sur lesquelles tous les mathématiciens sont d'accord. Mais cela n'est pas toujours aisé et je dois reconnaître que M. Le Dantec, en particulier, paraît se faire un point d'honneur de penser sur ces questions autrement que les mathématiciens ; il revient, dans un livre qu'il vient de publier, et qui est fort suggestif comme tout ce qu'il publie[2], sur une discussion que nous avons eue il y a quelques années et j'ai le regret de constater que mes arguments n'ont pas modifié ses conceptions. Je

1. La fin de ce chapitre est la reproduction d'un article paru dans la *Revue du Mois* du 10 juillet 1911 (t. XII, p. 77), sous le titre : *Les Probabilités et M. Le Dantec.*

2. *Le Chaos et l'Harmonie universelle* (Paris, F. Alcan, 1911).

voudrais essayer de les présenter avec plus de clarté : s'ils ne convainquent pas M. Le Dantec, peut-être intéresseront-ils quelque autre lecteur.

Je dois commencer par avouer qu'il est un point sur lequel M. Le Dantec est inattaquable : c'est lorsqu'il déclare préférer tel langage à tel autre. Il est parfaitement légitime, du moment que l'on prévient ses lecteurs, de dire *pessurku* au lieu de probabilité et *assurbé* au lieu d'espérance mathématique; le jour où tout le monde aurait adopté ces conventions, ce serait aussi clair.

Je comprends moins lorsque M. Le Dantec déclare que les mots : *probabilité d'un coup isolé* n'ont aucun sens. Si ces mots choquent M. Le Dantec, je veux bien les remplacer par une expression plus courte et dire, par exemple « *la fraction* » : c'est un nombre inférieur à l'unité que les mathématiciens savent calculer d'une manière précise lorsqu'on leur fait connaître les règles d'un jeu. Et si l'on donne le choix pour une partie isolée, entre deux jeux pour lesquels *les fractions* sont différentes, tout homme de bon sens et renseigné choisira le jeu pour lequel *la fraction* est la plus grande[1]. Il pourra, d'ailleurs,

1. Je suppose, bien entendu, que l'on est obligé de jouer l'une des parties et une seule et que l'on tient à gagner. Il est facile de construire des hypothèses où M. Le Dantec lui-même choisirait en tenant compte de *la fraction*; je lui ai signalé un de ces cas dans une lettre particulière et il en a convenu. M. Fréchet, de son côté, m'a écrit spontanément pour me communiquer un exemple de même nature.

J'avoue d'ailleurs que je m'explique mal l'intérêt de la remarque de M. Le Dantec, que le plus souvent les gros lots sont gagnés par des porteurs d'un seul billet. Si la remarque est exacte, elle prouve seulement que les porteurs d'un seul billet ont, dans leur ensemble, la majorité des billets.

après l'événement, regretter ce choix, mais ce regret ne l'empêchera pas d'agir suivant la même règle si l'occasion se représente. Telle est la valeur pratique universelle de *la fraction*, et cela suffit à condamner l'opinion que c'est un nombre vide de sens. A côté de cette valeur universelle, *la fraction* peut avoir pour tout homme qui n'est pas absolument hostile au jeu, une utilité bien plus large. Par exemple, peu de gens hésiteraient à dépenser un franc pour acquérir un billet unique d'une loterie de mille billets, pourvue d'un gros lot d'un million [1]. Mais nous sommes ici en dehors des mathématiques, car la psychologie de l'acheteur intervient, ainsi qu'une foule de conditions connues de lui seul. Si l'on propose de jouer en une seule partie de pile ou face, cent mille francs contre un million, l'offre a beau être théoriquement avantageuse, aucun mathématicien ne prétendra démontrer qu'il faut l'accepter, et je connais beaucoup de mathématiciens qui la refuseront, pour d'excellentes raisons.

18. — Mais le désaccord le plus grave entre M. Le Dantec et moi [2] est relatif à la succession d'un nombre indéfini de parties de pile ou face; je voudrais y insister un peu car M. Le Dantec me paraît faire ici une véritable erreur mathématique, et dès lors, je ne désespère pas de le convaincre ; la question est d'ailleurs assez déli-

1. Bien entendu, si l'on a le choix, on préférera acheter les mille billets ; nous supposons qu'un seul est à vendre.

2. On excusera, je pense, ce langage personnel qui est plus bref et, à tout prendre, plus modeste que celui par lequel je rappellerais à chaque instant que les idées qui sont les miennes appartiendraient aussi à tout mathématicien qui aurait réfléchi à ces questions.

cate, et insuffisamment connue; aussi n'est-il peut-être pas inutile d'entrer dans quelques détails.

Pour poser nettement la question, je demande la permission de reproduire le passage de mon livre qui a été l'origine de cette polémique[1].

Remarques sur quelques paradoxes. — Nous pourrions borner là l'exposé des principes essentiels de la théorie du jeu de pile ou face; ces principes étant bien établis, les conséquences que nous en déduirons par des raisonnements purement logiques sont rigoureusement démontrées, et par suite toute assertion contraire à ces conséquences devra être regardée comme inexacte, sans qu'il soit nécessaire d'examiner les arguments sur lesquels on prétend la baser. Cette manière de procéder est la plus conforme à l'esprit mathématique; il me paraît cependant préférable de ne pas m'y tenir, car tout le monde n'a pas l'esprit mathématique et, en ce qui concerne les questions de probabilité, beaucoup d'esprits, excellents par ailleurs, ont une certaine méfiance des raisonnements logiques et sont disposés à leur préférer des raisons de sentiment. J'ai eu récemment l'occasion de constater cette tendance chez un des esprits les plus distingués de notre temps, bien connu par ses publications scientifiques et philosophiques, et dont l'éducation mathématique a été très sérieuse. Il m'a dès lors semblé qu'il valait mieux ne pas traiter ces tendances par le pur dédain que serait en droit de leur opposer un mathématicien qui jugerait entièrement superflu de convaincre, du moment que ses raisonnements sont irréprochables. Il y a, en effet, à mon avis, un très grand intérêt scientifique et social à ce que les principes fondamentaux du calcul des probabilités soient admis sans restriction par le plus de personnes possible; si donc quelques arguments peuvent amener ce résultat, il vaut la peine d'y consacrer quelques lignes bien que, au point de vue mathématique absolu, ils soient inutiles.

L'une des principales sources de ces raisonnements paradoxaux sur lesquels nous aurons à revenir est la suivante :

1. *Eléments de la Théorie des probabilités* (première édition, Hermann, 1909; deuxième édition, Hermann, 1910), p. 18.

on considère un événement futur comme réalisé, sous prétexte que l'expérience a prouvé qu'il est extrêmement probable. On commet ainsi une erreur, sans doute très petite, mais l'accumulation répétée de telles erreurs suffit pour conduire à des conséquences entièrement inexactes. Nous étudierons plus loin le mécanisme mathématique de cette accumulation d'erreurs; pour l'instant, nous nous contenterons de montrer par l'absurde le défaut de ce genre de raisonnements.

Supposons que l'on joue un très grand nombre de parties successives à pile ou face et que l'on note tous les résultats ; on saura à chaque instant quel est le gain ou la perte d'un joueur qui aurait toujours parié pour pile, l'enjeu étant toujours le même à chaque partie. Si l'on fait effectivement l'expérience, on constate aisément qu'au bout d'un certain nombre de parties, souvent petit, dépassant rarement 100 et presque jamais 1.000, le gain et la perte se trouvent réduits à zéro ; lorsqu'on en est arrivé là, il y a une chance sur deux pour que la partie suivante amène un gain; s'il en est ainsi, nous dirons que l'ensemble des parties jouées constitue une bonne série; sinon, nous continuerons le jeu jusqu'à un nouveau retour à zéro et il y aura de nouveau une chance sur deux pour que la partie suivante soit gagnée; s'il en est ainsi, nous aurons une bonne série; sinon nous continuerons encore le jeu et nous finirons bien par arriver à obtenir *une bonne série*, puisque nous avons une chance sur deux d'y parvenir chaque fois que nous revenons au zéro. Pratiquement, l'expérience montrera à celui des lecteurs qui voudra la tenter que l'on obtient généralement une bonne série après un petit nombre de coups et que l'on y parvient *sûrement* [1] pourvu que l'on ait la patience de jouer, s'il est nécessaire, quelques milliers de coups.

Ceci étant admis, Paul joue avec Pierre à pile ou face et est décidé à prolonger le jeu jusqu'à ce qu'il ait réalisé une bonne série [2]; il réalise ainsi un gain égal à la mise. Il

1. Je souligne ce mot, car c'est par lui que s'introduit l'erreur de raisonnement ; il faudrait dire *presque sûrement*, c'est-à-dire que la probabilité pour ne pas réussir est extrêmement faible (voir au nº 18 du livre cité un calcul détaillé).

2. Si la première partie est gagnée, elle constitue à elle seule une bonne série.

cesse alors le jeu qu'il reprend le lendemain contre Jean; il le prolonge de même jusqu'à ce qu'il ait réalisé une bonne série; il peut ainsi continuer tous les jours et gagner régulièrement une somme égale à la mise. Si nous supposons maintenant que Paul joue toujours contre le même adversaire Pierre une suite indéfinie de parties, Paul peut, sans attendre au lendemain, considérer son jeu comme interrompu après chaque bonne série et recommencer à noter à ce moment-là; comme rien ne distingue un instant d'un autre et que l'on peut toujours supposer que le jeu commence à un instant quelconque, Paul réalisera ainsi un nombre illimité de bonnes séries successives et par suite, un gain illimité (ou, du moins, qui ne sera limité que par la lenteur du jeu et la durée de la vie humaine). Mais, sur la même succession de parties, Pierre peut faire le même raisonnement; son gain est donc aussi illimité, à condition que l'on puisse jouer assez longtemps; telle est la conséquence absurde à laquelle on aboutit : chacun des joueurs réalise un gain qui croît proportionnellement au temps [1]. Nous nous contenterons pour l'instant, d'avoir mis en garde le lecteur contre les inconvénients de certains modes de raisonnement; nous étudierons d'une manière plus approfondie cette question lorsque nous aurons acquis les principes nécessaires.

19. — Voici maintenant, si je ne me trompe, l'essentiel de l'objection de M. Le Dantec [2]:

Et l'auteur conclut : « Telle est la conséquence absurde à laquelle on aboutit : chacun des joueurs réalise un gain qui croît proportionnellement au temps. » Une note montre que l'auteur attribue cette absurdité à ce que l'on a considéré comme sûr, pratiquement, d'arriver à un moment où un joueur donné gagne un sou, *tandis que ce n'est que presque sûr*. Or ce n'est pas du tout là qu'est l'erreur;

1. Dans la deuxième édition (voir ses *errata*), j'ai, à la suite d'une correspondance avec M. Le Dantec, ajouté ici une phrase explicative qui, sans être indispensable, m'a paru rendre la pensée plus claire. « Ceci contredit absolument l'hypothèse initiale d'après laquelle quelques milliers de coups suffiraient *sûrement* pour rétablir l'équilibre entre les deux joueurs. »

2. *Loc. cit.*, pp. 160-161.

elle est, à mon avis, dans le raisonnement même de Borel. Si vous voulez vous reporter aux raisonnements de sens commun dont j'ai accompagné, dans le présent chapitre, la courbe figurative placée quelques pages plus haut, vous verrez bien qu'il y a, en effet, des moments où, au bout d'un assez grand nombre de coups, Pierre gagne n'importe combien et *d'autres moments* où Paul gagne aussi n'importe combien. *Mais ce ne sont pas les mêmes moments.* Si l'on arrête le jeu quand Pierre gagne n sous, Paul les perd au même moment; cela est de toute évidence, et il est néanmoins pratiquement certain que, de loin en loin, la courbe coupera l'axe des x, ce qui fait que, de temps en temps, les gains et les pertes s'équilibreront; mais entre les points où elle coupe l'axe des x, la courbe passe à certains moments par des *maxima* (gains de Pierre) et à *d'autres moments* par des *minima* (gains de Paul). J'ai fait remarquer cela à Borel; nous avons même eu à ce sujet une longue correspondance de laquelle est résulté un *erratum*, ajouté à la deuxième édition de la *Théorie des probabilités* (p. VII); or cet *erratum* prouve simplement que l'auteur a persisté dans sa manière de voir, et qu'il attribue l'erreur cachée sous un prétendu paradoxe à ce que l'on considère comme *sûr* un équilibre à venir qui n'est que *probable.*

Enfin, voici les « raisonnements de sens commun » auxquels M. Le Dantec fait allusion dans le passage précédent :

Convenons de représenter graphiquement, sur un papier quadrillé, l'état relatif des nombres de coups pile et de coups face au cours d'une longue série de coups. Nous prendrons comme axe des x un trait horizontal du quadrillage, et nous y porterons les nombres de coups. Marquons le zéro sur l'axe des x; ce sera le point de départ du jeu. Le premier coup amène pile; je le porte, sur l'ordonnée 1 en a; le second coup amène encore pile; je le porte sur l'ordonnée 2 en b. Le troisième amène face, je descends d'un échelon à partir de b et le porte en c sur l'ordonnée 3, et ainsi de suite. J'ai ainsi défini une courbe dont l'ordonnée correspondant à l'abscisse n représentera l'excès du nombre de coups pile sur le nombre de coups face au bout de

n coups. Nous savons d'avance quelle sera l'allure générale de la courbe. Ce sera une courbe sinueuse qui coupera de temps en temps l'axe des x mais sans qu'il y ait aucune régularité dans la distribution des points où elle le coupe. Nous pouvons prévoir quelques-unes des particularités de la courbe, en nous basant uniquement sur ce fait, que nous sommes certains que le jeu n'obéit à aucune loi.

La plus importante, et qui contient en réalité toutes les autres, est la suivante : Etant donné à l'avance un nombre N aussi grand qu'on le voudra, on doit prévoir qu'il arrivera un moment où l'ordonnée de la courbe sera égale à N, En d'autres termes, il n'y a pas de limites à l'écart qui se manifestera au cours d'un nombre très grand de parties entre le nombre des *pile* et le nombre des *face*.

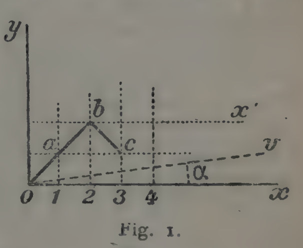


Fig. 1.

Cela découle immédiatement de la certitude dans laquelle nous nous trouvons relativement au fait que les coups déjà joués n'influent aucunement sur les coups à venir. Après deux coups, quand nous sommes au point b, nous pouvons prendre b comme origine et bx' comme axe de x. Alors l'allure de la courbe à partir de b devra être la même qu'à partir de o ; or l'écart à obtenir désormais est devenu (N — 2). Si à partir de b la courbe restait *toujours* au-dessous de l'axe bx' cela constituerait une loi; donc il y aura un moment où elle passera au-dessus et où l'écart à obtenir ne sera plus que (N — 3) ; nous reprendrons là une nouvelle origine, et avec les mêmes raisonnements nous verrons que s'impose la nécessité d'arriver à un moment où l'écart à obtenir ne sera plus que (N — N), c'est-à-dire où l'écart effectif obtenu à partir de o est N. Le théorème est démontré.

Il est évidemment vrai pour l'écart entre le nombre des face et le nombre des pile, comme pour l'écart entre le nombre des pile et le nombre des face. C'est-à-dire que la courbe sinueuse partie de o aura des ordonnées négatives qui ne le céderont en rien à ses ordonnées positives.

Une conséquence de ce théorème est que, si loin de l'axe des x que nous nous trouvions à un moment donné, nous

devons toujours nous attendre à y revenir au bout d'un nombre suffisant de parties ; en d'autres termes, il n'y a pas de moment à partir duquel la courbe ne coupera plus l'axe des x. Soit en effet N l'écart actuel : je le suppose positif. En prenant ce point de la courbe comme origine, nous sommes certains d'arriver un jour ou l'autre à l'écart (— N), c'est-à-dire à l'axe primitivement choisi comme axe des x.

Tout ceci est nécessaire parce que nous avons admis en commençant que le phénomène étudié ne saurait être soumis à aucune loi.

Une conséquence évidente de ces remarques est que, pour un nombre de coups très grand, le rapport de l'écart au nombre des parties jouées sera toujours très petit, quelle que soit d'ailleurs la valeur absolue de l'écart. En effet, puisque la série des zéros de notre courbe est illimitée, nous pourrons toujours jouer un nombre de coups assez grand pour que le nombre m_1, qui aboutit au dernier zéro avant le commencement de l'arc de courbe conduisant à l'écart, soit aussi grand que nous le voudrons. Or, à partir de zéro, l'écart n, dans les conditions les plus favorables, ne peut s'obtenir qu'au bout d'un nombre de coups au moins égal à n. Le nombre de coups conduisant à l'écart de n sera donc supérieur ou égal à $m_1 + n$, et le rapport de l'écart au nombre des coups sera inférieur à $\frac{n}{n + m_1}$, rapport que l'on pourra toujours rendre aussi petit qu'on le voudra en donnant à m_1 une valeur assez grande [1].

Le fait qu'aucune loi n'existe en faveur des coups pile par exemple, nécessite les rencontres successives de la courbe avec l'axe ox. Si la courbe, au lieu de prendre une allure vaguement symétrique par rapport à l'axe ox, prenait cette allure par rapport à un autre axe ov, faisant avec ox un angle α, aussi petit qu'on le voudra, il y aurait une loi en faveur des coups pile, ce qui est impossible, l'angle ne peut être que nul.

Voici des conséquences que des raisonnements de sens

1. Ce raisonnement « de sens commun » est incomplet, car il n'est pas du tout certain que l'écart n ne se produise pas avant m_1 coups. M. Le Dantec ne peut pas, malgré tout son talent, remplacer par des raisonnements vagues un calcul précis.

commun nous font tirer de la nécessité où nous sommes d'affirmer que le jeu de pile ou face n'obéit à aucune loi. La loi des grands nombres, c'est-à-dire la nullité de l'angle de notre figure, n'est donc que la transformation verbale de l'affirmation de l'absence de toute loi. Et le fait que des penseurs en ont conclu à l'existence d'une loi du hasard me rappelle cette plaisanterie d'une vieille opérette dans laquelle le chef des conjurés s'écrie, ne voyant pas apparaître à l'heure fixée le fanion attendu :

« L'absence de signaux serait-elle un signal? »

En réalité, il faut bien avouer que nous ne sommes pas très satisfaits de notre raisonnement par l'absurde. Si la loi des grands nombres ne se vérifiait pas au bout d'un nombre très grand de coups, nous en tirerions la démonstration d'une loi en faveur de pile ou de face, ce qui serait contraire à notre hypothèse; c'est bien là le raisonnement par l'absurde; ici comme partout où nous sommes réduits à l'employer à défaut d'un raisonnement positif, ce mode de démonstration nous déplaît. Nous pensons instinctivement à l'âne de Buridan. En réalité, notre raisonnement implique un postulat; nous savons bien qu'il ne saurait y avoir aucune loi dans le jeu de pile ou face, mais *notre jeu sera-t-il assez précis, assez souple pour nous démontrer, par lui-même, au bout d'un grand nombre de coups, cette absence remarquable de toute loi*? Nous n'oserions pas l'affirmer *a priori*. Nous n'oserions pas énoncer la loi des grands nombres si nous ne l'avions vérifiée expérimentalement un grand nombre de fois pour chaque jeu de hasard. Nos raisonnements *a priori* ne valent que par une démonstration *a posteriori*. Et d'ailleurs, notre loi des grands nombres n'a pas le caractère de quelque chose d'inévitable; ce n'est pas une loi en effet, puisque ce n'est que le résultat de l'absence de toute loi. Si nous constations expérimentalement un cas dans lequel, au bout d'un très grand nombre de coups, il y aurait toujours eu deux fois plus de coups pile que de coups face, nous n'aurions pas le droit de nous insurger; nous rechercherions d'abord s'il n'y a pas une erreur systématique dans le jeu; supposons qu'il n'y en ait pas; nous devrions en conclure qu'il n'y a pas de loi en faveur des coups pile, mais que notre jeu n'a pas, dans le cas considéré, réussi à mettre en évidence cette absence de loi. Je ne sache pas que cela soit arrivé jamais,

mais cela n'est pas *impossible*. Et cependant le théorème de Bernoulli donne une démonstration *a priori* de la loi des grands nombres! Cette loi serait donc vraiment une loi.

20. — J'ai tenu à citer textuellement ces raisonnements parce qu'ils sont un exemple aussi frappant que possible du fait que, dès que les problèmes de probabilités deviennent tant soit peu complexes, le bon sens, même servi par une intelligence claire et profonde, ne peut se passer de l'aide du calcul : il conduit tout au plus à des résultats non inexacts, mais incomplets et flous. Si M. Le Dantec avait examiné avec soin et repris personnellement les calculs auxquels je renvoie dans le passage cité plus haut (renvoi au n° 18 du livre cité), il est certain qu'il se serait rendu un compte plus précis de la marche générale d'une très longue partie de pile ou face. Je ne reproduirai pas ici ces calculs, préférant montrer comment on peut arriver à des résultats équivalents par une voie différente et peut-être plus intuitive.

Je rappellerai ce résultat bien connu que l'écart moyen à prévoir, au bout d'un certain nombre de parties de pile ou face, est grossièrement égal à la racine carrée de ce nombre de parties. Par exemple, au bout d'un million de parties, un écart de mille en plus ou en moins pour l'un des joueurs n'est pas exceptionnel. Plaçons-nous dans l'hypothèse où cet écart moyen est précisément réalisé et *demandons-nous quelle est la probabilité pour que l'écart soit toujours resté de même sens pendant le million de parties*[1]. Ce problème est précisément

1. Pour traiter cette question d'une manière rigoureuse et tout à fait générale, il faudrait envisager successivement toutes les hypothèses possibles après un million de parties (écart égal à zéro, à 1,

celui que M. Désiré André a étudié sous le nom de *problème du scrutin* et dont la solution est bien connue. Dans un scrutin uninominal, deux candidats A et B sont en présence; le premier A a m voix et le second B en a $m + n$; quelle est la probabilité pour que B conserve la majorité pendant toute la durée du dépouillement? La réponse est que la probabilité est égale au quotient de la majorité de B, c'est-à-dire de n par le total du nombre des bulletins, c'est-à-dire par $2m + n$. Si dans une partie de pile ou face, on sait seulement qu'au bout d'un million de parties, B a gagné 1.000 parties de plus que A, on peut dire qu'il y a une chance sur 1.000 pour que B ait toujours eu l'avantage.

D'une manière plus générale, le quotient d'un nombre par sa racine carrée, étant évidemment égal à cette racine carrée, la probabilité pour que l'un des joueurs conserve l'avantage pendant un grand nombre de parties est sensiblement égale à l'inverse de la racine carrée de ce nombre de parties[1].

En d'autres termes, si l'on envisage 100 millions de parties consécutives de pile ou face, il y a, en gros, une chance sur 10.000, pour que l'un

à 2, etc.) avec leurs probabilités respectives, faire pour chacune d'elles le calcul que nous faisons dans le cas de l'écart moyen, et sommer les résultats obtenus; ce calcul conduit à une intégration facile, dont le résultat est le même que celui du raisonnement abrégé, à un facteur numérique près, voisin de l'unité et sans influence sur nos conclusions.

1. Comme il a été dit dans la note précédente, ce résultat n'est exact qu'à un facteur numérique près, assez voisin de l'unité pour ne jouer aucun rôle dans la suite. Il faudrait diviser la probabilité par 2 si l'on fixait d'avance celui des deux joueurs qui doit conserver l'avantage; nous supposons ici qu'il nous est indifférent que ce soit l'un ou l'autre, pourvu que ce soit toujours le même.

des joueurs ne cesse pas de l'emporter, pour que la courbe de M. Le Dantec ne coupe pas l'axe des *x*. Supposons que les 2 millions d'habitants adultes de Paris, s'étant groupés deux par deux commencent demain matin à jouer à pile ou face en conservant chacun le même partenaire et en convenant de cesser le jeu lorsqu'ils seront « quitte ». En allant très vite, chaque couple de joueurs pourra peut-être jouer une partie par seconde, soit à peu près 10 millions de parties par an, avec la journée de huit heures. Eh bien, on doit prévoir qu'au bout de dix ans, il y aura encore une centaine de couples qui n'auront pas terminé, et qu'au bout de mille ans, si les joueurs ont confié leurs parties à leurs héritiers, une dizaine de parties continueront encore, l'un des deux partenaires ayant toujours été en déficit. On voit que nous sommes loin de la certitude pratique que l'équilibre se rétablit forcément : pour les joueurs qui sont morts à la peine, en courant après leur argent, tout s'est, au contraire, pratiquement passé comme si l'équilibre n'eût jamais dû se rétablir[1].

Je crois que ce résultat pourrait être utilement médité par les inventeurs de martingales : c'est, au fond, la croyance plus ou moins confuse à la certitude de l'équilibre qui est la source de leurs erreurs. Or, si l'on considère des tables de roulette *où le zéro serait supprimé*, le jeu de rouge ou noir (ou toute autre combinaison *simple*) est équivalent au jeu de pile ou face. On peut

1. Néanmoins, il reste légitime de dire que le rétablissement de l'équilibre est *certain*, si les joueurs ont assez de patience, car à mesure que les parties deviennent plus nombreuses, la probabilité pour qu'il n'y ait pas équilibre tend vers zéro.

considérer que chaque coup marque, pour le joueur qui arrive à cet instant, l'origine d'une partie nouvelle, de sorte que, dans une maison de jeu bien achalandée, possédant plusieurs roulettes, le nombre des origines est de plusieurs centaines de mille par an. Les diverses parties que l'on peut ainsi considérer isolément ne sont pas distinctes, puisqu'elles se superposent ; il est aisé de se rendre compte que ce défaut d'indépendance ne modifie que partiellement les conclusions tirées des parties de pile ou face indépendantes que nous imaginions tout à l'heure[1] ; il est certain que, si l'on considérait le tableau complet des résultats effectivement constatés l'année dernière, on constaterait qu'il y a eu, dans le courant de l'année, certains instants, très rares et très brefs, tels que le joueur entrant à ce moment-là dans la salle de jeu et jouant sans interruption toujours le même jeu (*rouge*, par exemple) à la même table, n'aurait pas cessé, pendant toute l'année, d'être en perte ; il n'aurait évidemment pas cessé d'être en gain s'il avait joué constamment le jeu contraire (*noire*) ; nous supposons, bien entendu, que le *zéro* est considéré comme inexistant, c'est-à-dire que les coups où il sort n'entraînent ni gain ni perte : nul n'ignore que la présence du zéro est indispensable aux bénéfices du tenancier, seule raison d'être des maisons de jeu.

21. — Mais il est temps de revenir au point précis sur lequel porte l'argumentation de M. Le

1. On pourrait d'ailleurs traiter le cas des parties non indépendantes par un raisonnement plus bref, où interviendraient les *maxima maximorum* et les *minima minimorum* de la courbe.

Dantec; les considérations précédentes contribuent, je crois, à éclaircir la question aux yeux du sens commun, en fournissant à ce sens commun une notion plus nette de l'allure générale du jeu, mais elles n'atteignent pas le fond même du débat. Il s'agit de savoir si la conclusion à laquelle conduisent certains raisonnements approximatifs, à savoir que chaque joueur réalise un gain qui croît *proportionnellement au temps*, est absurde ou non.

M. Le Dantec croit devoir faire observer que Paul et Pierre ne sont pas tous deux gagnants au même moment; je n'ose cependant croire que son opinion sur l'absence de sens commun chez les mathématiciens va jusqu'à lui faire supposer que je n'ai pas eu l'idée qu'ils pourraient sans absurdité être gagnants tous les deux, tour à tour, à des moments différents : un mathématicien qui serait assez absorbé dans ses calculs pour ne pas penser à une telle circonstance simple mériterait une place d'honneur dans le Panthéon des savants distraits. Ce qui est absurde, ce n'est pas que Pierre et Paul réalisent tous deux un gain à des moments différents, c'est que ce gain croisse proportionnellement au temps. M. Le Dantec trouvera les calculs et les chiffres précis au n° 18 de mon livre auquel renvoie le passage cité plus haut; je les résume ici : en un millier de parties de pile ou face, on peut espérer raisonnablement que l'équilibre entre les deux joueurs se rétablira une vingtaine de fois; comme, chaque fois que l'équilibre se rétablit, chaque joueur a une chance sur deux de gagner la partie suivante, chacun d'eux peut, en raisonnant comme plus haut, espérer se retirer après un gain d'environ

10 francs, si l'enjeu est de 1 franc. Ces espérances ne sont pas contradictoires : en fait, c'est celui qui se sera réservé le droit de cesser le jeu au moment où il lui convient qui sera le gagnant. Si cette espérance, au lieu d'être *presque toujours* réalisée, était *toujours* réalisée, le gain de 10 francs après 1.000 parties deviendrait un gain de 10.000 francs après un million de parties, de 10 millions après un milliard de parties. Or, ceci est absurde, pour bien des raisons, dont voici peut-être les deux les plus convaincantes : un écart de 10 millions sur un milliard de parties est infiniment peu vraisemblable d'après la loi des grands nombres, il est donc absurde de le supposer certain. De plus, si l'écart de 10 millions se produit successivement en faveur des deux joueurs, il faut évidemment qu'il soit joué au moins 10 millions de parties entre le moment où Paul gagne 10 millions et celui où l'équilibre se rétablit[1] ; par suite, il existe deux intersections successives de la courbe avec l'axe des x qui sont séparées par au moins 20 millions de parties; ceci contredit absolument l'hypothèse d'après laquelle l'équilibre se rétablirait *sûrement* au bout de quelques milliers de coups.

On voit que l'allure de la courbe construite par M. Le Dantec est plus complexe qu'il ne semble l'imaginer; les longues séries de coups avantageuses dans leur ensemble à l'un des joueurs sont assez peu probables; mais, lorsque le jeu dure assez longtemps, elles finissent tout de

1. Cette remarque de sens commun a été faite par M. Le Dantec, dans le passage cité plus haut : c'est un des cas qui sont assez simples pour que le sens commun puisse se passer de l'aide du calcul.

même par se produire, et elles peuvent alors empêcher l'équilibre de se produire pendant très longtemps ; par suite, la courbe commencera généralement par décrire de légères sinuosités autour de l'axe des x ; puis, il arrivera qu'elle s'en écartera assez notablement, et elle sera alors sinueuse, pendant quelque temps, non pas autour d'une droite inclinée (M. Le Dantec explique fort bien pourquoi ce serait absurde), mais autour d'une parallèle à l'axe des x ; et cela, jusqu'à ce qu'il se produise de nouveau un écart assez grand qui pourra, soit la ramener vers ox, soit l'en écarter encore davantage[1]. Ces indications sont grossières, mais il est difficile de les préciser, *car il n'y a pas de loi;* la seule affirmation un peu générale qu'on puisse énoncer, c'est que tout point de la courbe peut être pris comme origine et que, par suite, tout ce que l'on peut dire de l'allure par rapport à ox peut être dit aussi de l'allure par rapport à une parallèle quelconque à ox que la courbe aura coupée une fois ; au bout d'un temps assez long, *toute parallèle* à ox sera coupée une fois[2] ; on ne peut donc rien dire sur ox qui ne pourrait être dit d'une autre parallèle quelconque à ox, du moment que l'on étudie le jeu pendant assez longtemps ; c'est seulement pen-

1. M. Le Dantec remarque bien ce rôle des parallèles à l'axe des x, mais ne tire pas de cette remarque toutes les conséquences qu'elle comporte, sinon, il aurait observé que, toutes les parallèles à l'axe des x jouant le même rôle, si l'axe des x est coupé par la courbe *sûrement* à des intervalles assez fréquents, il doit en être de même de toutes les autres parallèles, ce qui est évidemment absurde.

2. Plus généralement, si l'on suppose le jeu prolongé indéfiniment dans les deux sens, toute parallèle à ox est coupée un nombre de fois proportionnel à la racine carrée de l'intervalle de temps considéré, du moment que cet intervalle est assez grand.

dant les quelques instants qui suivent l'instant actuel que l'on peut tirer quelque conclusion particulière à *ox* du fait que c'est *l'origine à l'instant actuel*. Ces remarques s'appliquent à toutes les questions de probabilités ; elles ne sont pas, je crois, sans relation avec ce que les physiciens modernes appellent le principe de relativité.

22. — Ce n'est pas seulement sur ce point particulier que M. Le Dantec prend à parti les mathématiciens ; il insiste beaucoup, à diverses reprises, sur l'impossibilité d'arriver par un raisonnement abstrait à la connaissance d'une loi naturelle. Je crois que les mathématiciens sont tous d'accord avec M. Le Dantec sur ce point ; ils ont parfois même poussé trop loin la défiance vis-à-vis des conséquences que l'on peut tirer du calcul ; c'est ainsi que Joseph Bertrand, en critiquant Maxwell, est allé jusqu'à nier la possibilité de la théorie cinétique, dont les promoteurs ont depuis, à l'exemple du Grec, prouvé le mouvement en marchant et en faisant même un assez beau chemin. Les mathématiciens, n'en déplaise à M. Le Dantec, distinguent en général fort bien, dans leurs écrits, ce qui est déduction pure de ce qui est raison de sentiment et opinion et, autant ils sont intransigeants lorsqu'on conteste leurs calculs, autant ils sont prêts à admettre que toute personne de bon sens, ayant bien compris les résultats de ces calculs, est libre de les interpréter pratiquement suivant sa psychologie personnelle[1]. Seulement, il est nécessaire, pour raison-

1. Voir plus loin, ch. VIII.

ner sur les probabilités, d'avoir compris par quel mécanisme analytique simple et inattaquable, la notion de l'absence de toute loi conduit à des conclusions, nécessairement toujours incertaines, mais dont l'incertitude peut être mesurée avec rigueur et souvent évaluée à une fraction dont la petitesse confond l'imagination. On revient ainsi, en dernière analyse, à la notion de probabilité d'un coup isolé que conteste M. Le Dantec ; car il s'agit de savoir si telle combinaison, dont la probabilité a été évaluée, par exemple, à un millionième de millionième, se produira ou non. La seule règle pratique d'action est d'agir comme si elle ne devait pas se produire. Si la probabilité est supposée encore plus faible, est par exemple 10^{-100}, on devra considérer qu'en la regardant comme nulle, on ne sacrifie en rien la rigueur scientifique, car il n'est pas de connaissance humaine qui soit vraie avec cette précision. C'est en ce sens, et en ce sens seulement, que des calculs purement abstraits peuvent conduire à des règles d'action concrètes, en permettant d'interpréter avec plus de précision ce que nous savons de la réalité ; c'est en ce sens que l'absence de toute loi est souvent une loi. Lorsqu'on a bien compris ce fait essentiel, il importe peu que le calcul de Bernoulli soit qualifié de stratagème : il permet, dans certains cas, de *prévoir* avec certitude (j'entends avec une certitude qui n'est jamais dépassée dans les prévisions humaines).



CHAPITRE III

PROBABILITÉS DISCONTINUES ET PROBABILITÉS CONTINUES

23. — On donne le nom de probabilités discontinues aux questions de probabilités dans lesquelles les cas possibles sont en nombre fini, par opposition avec les problèmes de probabilités continues [1], dans lesquels il y a une infinité de cas possibles (les cas possibles étant, par exemple, les diverses positions d'un point sur une droite).

Le jeu de pile ou face est le cas le plus simple des probabilités discontinues, puisque le nombre des cas possibles est égal à 2, et que leurs probabilités sont égales. Les problèmes que l'on peut se poser à propos des jeux de hasard, les dés, les cartes, appartiennent à la catégorie des probabi-

1. On peut envisager aussi les *probabilités dénombrables*, cas intermédiaire entre les probabilités continues et discontinues (voir ch. VII).

lités discontinues. Ce sont ces problèmes sur les jeux de hasard qui ont donné lieu aux premières recherches sur les probabilités ; sans l'intérêt porté à ces problèmes de jeux par des hommes comme Pascal, il est vraisemblable que la théorie des probabilités n'aurait été créée que bien plus tard : elle ne se serait pas trouvée en quelque sorte toute prête lorsqu'on en a eu besoin pour des applications pratiques ou scientifiques [1]. Ceux de nos contemporains qui emploient une fraction notable de leur temps à jouer au baccara ou au bridge auraient tort d'en conclure qu'ils font une besogne utile à la science ; la théorie des probabilités n'a plus besoin des jeux de hasard. Mais peut-être pourrait-on, de l'histoire de ses origines, retirer cet enseignement, qu'il est difficile de prévoir les répercussions scientifiques de spéculations en apparence futiles, si ces spéculations contribuent à aiguiser l'esprit et à créer des formes nouvelles de raisonnement.

Nous ne reviendrons pas sur la définition de la probabilité, définition qui, dans le cas des probabilités discontinues, revient à celle de *l'égale probabilité des cas*. Cette égale probabilité sera postulée tout au moins comme un cas limite.

24. — Désignons par N le nombre total des cas possibles (supposés tous également probables,

1. L'un des problèmes posés par le jeu qui a particulièrement contribué à exercer la sagacité des fondateurs de la théorie des probabilités est le *problème des partis* : deux joueurs ont convenu que les enjeux appartiendront à celui d'entre eux qui, par exemple, aura pu inscrire le premier à son actif dix parties gagnées ; ils interrompent le jeu, l'un ayant gagné 7 parties et l'autre 8 ; quels sont alors les *partis* des deux joueurs, c'est-à-dire comment les enjeux doivent-ils être partagés entre eux ?

bien entendu : nous ne reviendrons pas sur cette restriction) et par n le nombre des cas favorables ; la probabilité p est donnée par la formule

$$p = \frac{n}{N};$$

on voit que la probabilité est une fraction *toujours inférieure à l'unité ;* elle devient égale à l'unité dans le cas seulement où $n=N$: tous les cas sont favorables ; la probabilité se transforme en *certitude*.

En même temps que la probabilité d'un événement désiré, on a souvent intérêt à considérer la probabilité de l'événement contraire (nous supposons, pour l'instant, que tous les événements autres que l'événement favorable sont groupés sous cette même dénomination). Par exemple une urne renferme N boules parmi lesquelles n sont blanches, les autres étant rouges, noires, etc. On regarde comme l'événement favorable l'extraction d'une boule blanche ; l'événement contraire (ou défavorable) est alors l'extraction d'une boule qui ne soit pas blanche ; si nous désignons sa probabilité par q, nous avons

$$q = \frac{N-n}{N} = 1 - \frac{n}{N} = 1 - p$$

car le nombre des cas favorables à cet événement contraire est évidemment $N-n$, puisqu'il y a $N-n$ boules qui ne sont pas blanches.

La formule peut s'écrire

$$p + q = 1$$

c'est-à-dire que : *la somme des probabilités d'un événement et de l'événement contraire*

est égale à l'unité. Il importe de bien remarquer que l'on entend ici par événement contraire tout ce qui n'est pas l'événement regardé comme favorable, on n'admet pas en ligne de compte diverses alternatives : par exemple, s'il s'agit pour un joueur de gagner une partie d'échecs, l'événement contraire comprend tous les cas où il ne la gagne pas : partie perdue et partie nulle.

Lorsqu'un événement est très probable, sa probabilité p est très voisine de l'unité et la probabilité q est par suite très voisine de zéro. On dit quelquefois, dans le langage ordinaire, que la probabilité est *très grande ;* cette manière de parler ne s'accorde pas avec la définition mathématique d'après laquelle p ne doit pas dépasser l'unité : au lieu de dire *probabilité très grande* on devrait dire *probabilité très voisine de un.* Le langage usuel a son origine dans le fait suivant : en réalité, on compare l'éventualité de l'événement favorable à celle de l'événement contraire et lorsque la première éventualité dépasse de beaucoup la seconde, on dit que la probabilité est très grande : ce qui est très grand, *c'est le rapport de la probabilité de l'événement considéré à la probabilité contraire.* Ce rapport pourrait être appelé *probabilité relative* des deux événements, alors que la probabilité définie plus haut serait nommée la *probabilité absolue* ; mais nous ne modifierons pas le langage usuel et dirons simplement *probabilité* et non *probabilité absolue*, car c'est cette probabilité qui intervient presque toujours et il vaut mieux ne pas allonger inutilement l'expression qui la désigne. Il est toutefois bon de savoir que,

dans le langage ordinaire, non scientifique, on emploie souvent le mot probabilité dans le sens de probabilité relative, c'est-à-dire que l'on pense au rapport R défini par la formule

$$R = \frac{p}{q} = \frac{p}{1-p} = \frac{n}{N-n}.$$

Par exemple, si une urne renferme 1 000 boules, dont 999 sont blanches et 1 seule noire, on a en regardant comme favorable l'extraction d'une boule blanche

$$p = 0,999$$
$$q = 0,001$$
$$R = 999.$$

La probabilité p de l'événement favorable est très voisine de 1 ; sa probabilité relative R est très grande.

Lorsque l'on envisage à la fois plusieurs éventualités qui s'excluent réciproquement et dont l'ensemble embrasse tous les cas possibles, la somme de leurs probabilités est égale à 1. Supposons, par exemple, qu'une urne renferme N boules identiques, sauf que chacune porte inscrit l'un des nombres 1, 2, 3, ..., m ; on désignera par n_1 le nombre des boules portant le nombre 1, par n_2 le nombre de celles qui portent le nombre 2, etc. La probabilité p_1 d'extraire l'une des boules qui portent le nombre 1 est évidemment

$$p_1 = \frac{n_1}{N}.$$

On a de même :

$$p_2 = \frac{n_2}{N}, \ldots\ldots, p_m = \frac{n_m}{N}$$

et par suite :

$$p_1 + p_2 + \ldots + p_m = \frac{n_1 + n_2 + \ldots + n_m}{N} = 1$$

car la somme $n_1 + n_2 + \ldots + n_m$ est égale au nombre total N des boules.

25. — Considérons une urne renfermant N boules, parmi lesquelles a sont rouges, b sont blanches, les autres n'étant ni blanches ni rouges. La probabilité α d'extraire une boule rouge est

$$\alpha = \frac{a}{N}.$$

La probabilité β d'extraire une boule blanche est

$$\beta = \frac{b}{N}.$$

La probabilité p d'extraire une blanche *ou* rouge est évidemment

$$p = \frac{a+b}{N}$$

et l'on aperçoit immédiatement la relation

$$p = \alpha + \beta$$

qui exprime le *théorème des probabilités totales*. Pour énoncer ce théorème, il importe de bien préciser les conditions dans lesquelles nous avons obtenu la relation précédente. Nous considérons comme événement favorable, dont nous recherchons la probabilité p, l'extraction d'une boule rouge *ou* blanche ; cet événement favorable peut se produire de deux manières différentes, qui s'excluent réciproquement, car si la boule qu'on extrait est rouge, elle n'est pas blanche.

Les probabilités respectives de ces deux éventualités sont α et β ; *la probabilité p est égale à leur somme.* Il est visible que le raisonnement est le même dans le cas où le nombre des éventualités favorables est supérieur à deux ; si, par exemple, on considère comme favorable l'extraction d'une boule rouge, ou blanche, ou verte, ou violette, ou jaune, les autres couleurs étant exclues. On peut donc énoncer sous la forme suivante le théorème des probabilités totales.

THÉORÈME. — *Lorsque l'événement dont on recherche la probabilité peut se produire de plusieurs manières différentes qui s'excluent réciproquement, la probabilité cherchée est égale à la somme des probabilités partielles correspondant à ces diverses manières.*

Dans les applications de ce théorème, il est très important de vérifier que la condition d'*exclusion réciproque* est bien vérifiée.

26. — Considérons deux urnes dont la première renferme N boules, parmi lesquelles a sont blanches et dont la seconde renferme N′ boules, parmi lesquelles a' sont blanches. La probabilité α d'extraire une boule blanche de la première urne est

$$\alpha = \frac{a}{N}$$

et la probabilité α' d'extraire une boule blanche de la seconde urne est

$$\alpha' = \frac{a'}{N'}.$$

Supposons maintenant que l'on extraie une boule

de chaque urne. Quelle est la probabilité p pour que ces boules soient blanches toutes les deux ? Évaluons le nombre des cas possibles ; pour cela supposons que nous ayons numéroté, afin de les distinguer entre elles, les boules de chaque urne ; celles de la première sont numérotées de 1 à N et celles de la seconde de 1 à N'. L'extraction peut amener la boule 1 de la première urne et l'une quelconque des N' boules de la seconde ; cela fait N' cas, tous également probables ; on peut aussi amener la boule 2 de la première urne et l'une quelconque des N' boules de la seconde ; cela fait encore N' cas, également probables entre eux, et de même probabilité que les précédents ; à chacune des N boules de la première urne correspondent ainsi N' cas ; il y a donc en tout NN' cas possibles [1]. On compterait de même le nombre des cas favorables ; à chacune des a boules blanches de la première urne correspondent a' cas favorables, correspondant respectivement à l'extraction de chacune des a' boules blanches de

1. Si nous supposons : $N = 6$ et les N boules numérotées 0, 1, 2, 3, 4, 5 ; $N' = 10$ et les N' boules numérotées 0, 1, 2, 3, 4, 5, 6. 7, 8, 9, les résultats des extractions pourront être représentés par le tableau suivant, où l'on a inscrit d'abord le numéro de la boule sortie de la première urne :

00	10	20	30	40	50
01	11	21	31	41	51
02	12	22	32	42	52
03	13	23	33	43	53
04	14	24	34	44	54
05	15	25	35	45	55
06	16	26	36	46	56
07	17	27	37	47	57
08	18	28	38	48	58
09	19	29	39	49	59

Il est visible que les cas possibles sont au nombre de 60.

la seconde urne ; il y a donc aa' cas favorables et la probabilité p est donnée par la formule

$$p = \frac{aa'}{NN'}.$$

On voit que l'on a

$$p = \alpha\beta$$

et cette formule exprime le *théorème des probabilités composées*. Avant d'énoncer ce théorème, nous généraliserons les conditions dans lesquelles nous l'avons établi ; supposons que nous ayons trois urnes, dont l'une sera toujours appelée *la première* et dont les deux autres, supposées peintes extérieurement, l'une en blanc, l'autre en noir, s'appelleront la *seconde urne blanche* et la *seconde urne noire*. On extrait d'abord une première boule de la première urne ; si cette première boule extraite est blanche, on extrait la seconde boule de la *seconde urne blanche* ; si la première boule extraite n'est pas blanche, on extrait la seconde boule de la *seconde urne noire*. Quelle est la probabilité pour avoir deux boules blanches ? Il est clair que si l'on désigne par α la probabilité d'extraire une boule blanche de la première urne et par β la probabilité d'extraire une boule blanche de la seconde urne *blanche*, on a encore

$$p = \alpha\beta.$$

En effet, on ne se sert de la seconde urne *noire* que dans le cas où le premier tirage n'a pas donné une boule blanche ; donc, dans ce cas on ne se trouve certainement pas dans un cas favorable et le résultat du second tirage est sans importance ;

rien ne serait donc changé à la probabilité des cas favorables si l'on n'effectuait pas ce second tirage, ou bien si on l'effectuait dans une urne de composition absolument quelconque, *par exemple dans la seconde urne blanche ;* nous nous trouvons ainsi ramenés au cas déjà traité, mais ces remarques permettent de donner un énoncé plus général.

THÉORÈME. — *Lorsque l'événement dont on recherche la probabilité consiste dans la production successive de deux événements, la probabilité cherchée est égale au produit de la probabilité du premier de ces événements par la probabilité pour que le second se produise lorsque le premier s'est produit. Plus généralement, si la production successive de plusieurs événements est nécessaire, on doit multiplier les probabilités diverses de ces événements, en évaluant chacune d'elles dans l'hypothèse où l'on sait que les précédents se sont produits.*

Nous bornant au cas de deux événements, il est clair que, si la probabilité du second ne varie pas *lorsque le premier s'est produit*, cette dernière restriction est inutile ; l'on peut, dans ce cas, considérer les événements comme simultanés.

27. — Dans bien des cas, au lieu d'utiliser les théorèmes des probabilités totales ou composées, il est plus simple de recourir à l'énumération complète des cas, par les procédés de l'analyse combinatoire. En voici un exemple :

Quelle est la probabilité d'avoir d'entrée trois atouts à l'écarté? Rappelons que l'écarté

se joue avec 32 cartes, dont l'une se trouve retournée et indique l'atout ; chaque joueur reçoit 5 cartes. Les 5 cartes sont donc choisies au hasard parmi les 31 cartes autres que la retourne, dont 7 sont des atouts. Le nombre de jeux possibles, lorsque l'atout est fixé, est donc égal au nombre de manières dont on peut choisir 5 cartes parmi 31 cartes, c'est-à-dire à

$$\frac{31.30.29.28.27}{1.2.3.4.5}$$

c'est le nombre des cas possibles.

Parmi ces jeux, on obtiendra le nombre de ceux qui renferment 3 atouts en multipliant le nombre des manières de choisir 3 atouts parmi les 7 atouts par le nombre de manières de choisir 2 cartes quelconques parmi les 24 cartes qui ne sont pas des atouts ; on obtient ainsi pour le nombre des cas favorables

$$\frac{7.6.5}{1.2.3} \cdot \frac{24.23}{1.2}.$$

La probabilité demandée est égale au quotient du nombre de cas favorables par le nombre des cas possibles ; c'est-à-dire

$$\frac{7.6.5.24.23}{31.30.29.28.27} \cdot \frac{1.2.3.4.5}{1.2.3.1.2}.$$

Des réductions immédiates donnent ($6 \times 5 = 30$; $7 \times 4 = 28$) :

$$\frac{24.22.5}{31.29.27.2} = \frac{4.23.5}{31.29.9} = \frac{460}{8\,091} = \frac{1}{17{,}5\ldots}.$$

La probabilité demandée est comprise entre $\frac{1}{17}$ et $\frac{1}{18}$.

28. — Dans beaucoup de problemes sur les probabilités discontinues, il est commode d'introduire la notion d'*espérancc mathématique*, qui prête à des calculs généralement plus simples. Cette expression : *espérance mathématique* a été souvent critiquée et il faut reconnaître qu'elle n'est pas des mieux choisies ; mais elle est consacrée par l'usage et d'autre part, son emploi ne peut prêter à aucune confusion si l'on s'habitue à la considérer comme un terme unique, dont le sens ne résulte pas de la juxtaposition des sens des deux composants [1]. Si l'on voulait cependant chercher à interpréter ainsi le sens, il semble qu'il faudrait prendre *mathématique* dans le sens de *numérique* et traduire *espérance mathématique* par *valeur numérique de l'espérance*. L'espérance mathématique est en effet la valeur numérique que l'on est conduit forcément à attribuer à l'espérance d'un joueur, si l'on s'impose d'évaluer cette espérance numériquement. Car cette évaluation doit évidemment satisfaire aux conditions suivantes : dans un jeu tel que pile ou face, où les probabilités de gain des deux joueurs sont égales, leurs espérances mathématiques sont aussi égales ; de plus, la somme de ces deux espérances est égale à l'enjeu total, car on peut concevoir que les deux joueurs se mettent d'accord pour se partager cet enjeu au lieu de jouer. Ces simples remarques suffisent pour évaluer l'espérance mathématique dans les cas simples.

1. C'est le cas pour beaucoup de mots composés ; sans citer les termes de la langue vulgaire, tels que *chèvrefeuille*, *vinaigre*, etc., on en trouve des exemples fréquents dans le langage scientifique ; les mots *pile électrique* n'évoquent plus l'image de la pile de Volta ; de même, en mécanique, *force vive*, etc.

Par exemple, si je dois toucher 2 francs au cas où une partie de pile ou face donnera pile, mon espérance mathématique est 1 franc et peut être vendue ce prix à un joueur disposé à prendre ma place. Dans ce cas-là, je puis espérer trouver facilement à *vendre* mon espérance mathématique ; il n'en serait pas de même si l'enjeu de la partie de pile ou face était de cent millions ; dans certains cas, la valeur *commerciale* d'une espérance mathématique est donc inférieure à sa valeur numérique ; dans certains cas, au contraire, la valeur commerciale de l'espérance mathématique est supérieure à sa valeur numérique : c'est une conséquence du goût du public pour les loteries et des restrictions mises par la loi à leur développement ; un billet de loterie auquel est attaché une espérance mathématique de trente centimes [1] trouve facilement preneur à un franc.

Le grand avantage de la notion de l'espérance mathématique est le suivant : la combinaison de probabilités diverses entraîne à des calculs parfois compliqués ; au contraire, pour les espérances mathématiques, la règle est simple et intuitive : *pour avoir l'espérance mathématique totale qui s'attache à plusieurs événements fortuits, il suffit de faire la somme des espérances mathématiques correspondant à chacun d'eux.*

Paul doit jouer trois parties de pile ou face ; d'après les conventions faites, s'il gagne la pre-

1. Pour trouver l'espérance mathématique attachée à un billet de loterie, on n'a qu'à diviser la somme totale des lots par le nombre des billets ; ce serait la valeur que devrait avoir le billet si l'entrepreneur de la loterie ne faisait aucun gain et n'avait pas de frais de publicité.

mière, il touchera 10 francs ; s'il gagne la seconde 20 francs ; s'il gagne la troisième, 40 francs. Son espérance mathématique totale est 5 + 10 + 20 = 35 francs ; telle est la somme pour laquelle il pourrait vendre ses diverses chances à un joueur disposé à jouer un jeu équitable. Si ce fait ne paraît pas évident, il suffit d'observer que si un même acheteur achète l'espérance mathématique de Paul et celle de son partenaire, il doit les payer toutes deux le même prix, puisque la situation de Paul et celle de son partenaire sont identiques ; or cet acheteur est assuré de toucher 70 francs, quoi qu'il arrive ; donc, pour que le jeu ne lui apporte ni bénéfice assuré, ni perte certaine, il doit payer 35 francs à chacun des deux partenaires.

On doit admettre comme évident que si un jeu se compose de plusieurs parties succesives et si chaque partie est équitable, le jeu est équitable dans son ensemble ; ce principe est d'une grande utilité pour la solution de certaines questions.

29. — Appliquons la notion d'espérance mathématique à la solution du problème suivant qu'on appelle *problème de la poule :*

Trois joueurs A, B, C, *jouent à pile ou face dans les conditions suivantes :* A *et* B *font une première partie, à la suite de laquelle le perdant se retire et cède sa place à* C ; *il en est de même après chaque partie : le perdant se retire pour céder sa place au troisième joueur ; le joueur qui aura gagné deux parties consécutives touchera une somme* m ; *quelle est l'espérance mathématique de chaque joueur : 1° après la première partie ? 2° au début du jeu ?*

Supposons que A gagne la première partie et désignons après cette première partie, par a, b, c, les espérances mathématiques de A, B, C qui sont respectivement le joueur qui reste, celui qui sort et celui qui rentre. Si A gagne la seconde partie, le jeu sera fini et A en sera le gagnant ; si A la perd, il deviendra le joueur qui sort, B sera celui qui rentre et C celui qui reste. Dans la première hypothèse (A gagnant), A touche m, B et C ne touchent rien ; dans la seconde (A perdant) l'espérance mathématique de A devient b, celle de B devient c, et celle de C devient a. Nous écrirons que l'espérance mathématique de chacun des joueurs est égale à la somme des espérances mathématiques qui résultent de chacune de ces deux hypothèses ; comme la probabilité de chaque hypothèse est *un demi* nous obtenons ainsi les équations

$$a = \frac{m}{2} + \frac{b}{2}$$

$$b = 0 + \frac{c}{2}$$

$$c = 0 + \frac{a}{2}$$

qui sont faciles à résoudre et donnent

$$a = \frac{4}{7}\,m\,; \qquad b = \frac{1}{7}\,m\,; \qquad c = \frac{2}{7}\,m.$$

Telles sont les espérances mathématiques après la première partie ; pour obtenir les espérances mathématiques a', b', c' avant cette première partie, il suffit d'observer que l'espérance mathématique de C ne peut pas être modifiée par le résultat de cette première partie ; d'autre part, *la*

somme des espérances mathématiques de A et de B n'est pas non plus modifiée ; car, pour celui qui aurait acheté à A et à B leurs chances de gain, il est indifférent que l'un ou l'autre gagne cette première partie ; enfin, on a forcément $a' = b'$; on obtient ainsi finalement

$$a' = b' = \frac{a + b}{2} = \frac{5}{14} m \qquad c' = c = \frac{2}{7} m = \frac{4}{14} m.$$

On voit qu'au début du jeu, la situation des joueurs A et B est un peu meilleure que celle du joueur C ; si la somme à gagner m est 14 francs, l'espérance mathématique de A ou de B vaut 5 francs, et celle de C vaut seulement 4 francs.

La raison pour laquelle la notion d'espérance mathématique est plus maniable que la notion de probabilité, c'est que les probabilités ne peuvent pas s'ajouter par la même règle simple que les espérances mathématiques, en raison de la complexité des cas possibles ; l'addition des probabilités conduirait d'ailleurs à des probabilités supérieures à l'unité, ce qui serait absurde.

Considérons par exemple un joueur de dés qui doit toucher 6 francs chaque fois que le dé qu'il jette amène le point 5 ; son espérance mathématique par coup est de 1 franc ; sa probabilité de gagner un coup isolé est $\frac{1}{6}$. S'il joue 6 coups, son espérance mathématique est 6 francs, c'est-à-dire est égale à l'enjeu ; mais il serait absurde de croire que sa probabilité de gagner est égale à 6 fois $\frac{1}{6}$, c'est-à-dire à l'unité ; il peut fort bien arriver qu'aucun des 6 coups ne lui soit

favorable ; mais il peut arriver aussi qu'il gagne *plusieurs fois* et c'est grâce à cette possibilité de gagner plus de 6 francs que son espérance mathématique peut être égale à 6 francs, bien que dans certains cas il ne gagne rien. En réalité, la probabilité pour qu'il gagne 6 parties et touche 36 francs est $\left(\frac{1}{6}\right)^6$; la probabilité pour qu'il en gagne 5 et touche 30 francs est $6\left(\frac{1}{6}\right)^5 \frac{5}{6}$,..... la probabilité pour qu'il en gagne une seule et touche 6 francs est $\left(\frac{5}{6}\right)^5$; on pourrait obtenir l'espérance mathématique totale en multipliant chaque gain possible par sa probabilité et faisant la somme des résultats ; ce calcul, qui peut être simplifié par l'utilisation de la formule du binome, conduit au même résultat que le calcul direct et élémentaire : l'espérance mathématique est 6 francs.

Lorsque les probabilités sont très petites par rapport au nombre des épreuves et que la répétition de l'événement favorable peut être négligée, l'espérance mathématique fournit une approximation de la probabilité. Par exemple, Pierre possède 2 billets d'une loterie qui comporte 10.000 billets et 10 lots de 1.000 francs ; son espérance mathématique est 2 francs ; on en déduit que la probabilité pour qu'il gagne 1.000 francs est approximativement le quotient de 2 par 1.000, c'est-à-dire 0,002 ; en fait, la probabilité de gagner 1.000 francs est égale au produit de 0,002 par la fraction $\frac{1110}{1111}$, c'est-à-dire 0,001998... et la probabilité de gagner 2.000 francs est égale à $\frac{1}{1000} \times \frac{1}{1111}$.

Nous étudierons plus loin des phénomènes physiques (radioactivité) pour lesquels l'espérance mathématique a une expression bien plus simple que la probabilité.

30. — Pour terminer ce qui a trait aux probabilités discontinues, indiquons comment l'on doit calculer l'*unité décimale* d'écart pour pouvoir appliquer la règle simple que nous avons exposée en détail dans le cas du jeu de pile ou face.

Désignons par p la probabilité de l'événement observé et par q la probabilité complémentaire ($q = 1 - p$). *L'unité décimale d'écart s'obtient en multipliant la racine carrée du nombre des expériences par* $2\sqrt{pq}$. Considérons, par exemple 6.000 coups joués avec un seul dé ; la probabilité d'amener le point 3, par exemple, est $\frac{1}{6}$; la probabilité complémentaire est $\frac{5}{6}$; l'unité décimale d'écart est donc

$$2\sqrt{\frac{1}{6} \cdot \frac{5}{6}}\sqrt{6.000} = 57,7.$$

Sur 6.000 coups, on doit s'attendre à voir sortir 1.000 fois le point 3 ; la probabilité d'un écart égal ou supérieur à n fois l'unité décimale est, suivant la règle générale, égale à 10^{-n^2}. La probabilité d'un écart supérieur à 58 est donc un dixième, la probabilité d'un écart supérieur à 116 est un dix-millième, la probabilité d'un écart supérieur à 174 est un milliardième[1].

1. La formule approchée que nous employons ne fait pas de distinction entre les écarts positifs et négatifs ; il est évident qu'un écart dépassant 1 000 est possible s'il est positif et impossible s'il est négatif, car il est possible d'amener le point 3 plus de 2 000 fois et

Lorsque la probabilité p est très faible, la probabilité q est très voisine de l'unité ; on peut alors, au degré d'approximation que nous désirons, remplacer q par l'unité, on remarquera alors que le produit de la racine carrée de p par la racine carrée du nombre des expériences est égal à la racine carrée du nombre attendu de cas favorables[1] ; on peut donc énoncer la règle encore plus simple :

Dans le cas où la probabilité est faible[2], *l'unité décimale d'écart est le double de la racine carrée du nombre attendu des événements favorables*[3]. *La probabilité d'un écart dépassant n fois cette unité est* 10^{-n^2}.

Supposons, par exemple, que l'événement attendu soit la sortie d'un numéro déterminé au *lotto*, comportant 90 numéros, la probabilité est $\frac{1}{90}$; sur 9.000 expériences, on pourra s'attendre à 100 succès ; l'unité décimale d'écart est le double de la racine carrée de 100, c'est-à-dire 20 ; il y a donc une probabilité égale à *un dixième* pour que l'écart dépasse 20, c'est-à-dire pour que le numéro choisi sorte moins de 80 fois ou plus de 120 fois[4]. La probabilité d'un écart 40 sera un

il n'est pas possible de l'amener moins de zéro fois. Mais la probabilité de l'amener plus de 2 000 fois est si faible qu'on peut pratiquement négliger cette dissymétrie.

1. Car on a $\sqrt{p}\sqrt{n} = \sqrt{np}$.

2. Inférieure à un dixième, pour fixer les idées.

3. Dans le cas du jeu de pile ou face, l'unité décimale d'écart est égale au produit par la racine carrée de 2 du nombre des événements attendus ; on doit donc s'attendre en moyenne à des écarts $\sqrt{2}$ fois plus grands lorsque l'événement attendu a une probabilité faible.

4. La probabilité d'un écart aussi grand sur 200 parties de pile ou face, pour lesquelles on s'attend à 100 coups pile, est *un centième* au lieu d'un dixième.

dix-millième ; si l'entrepreneur du *lotto* pense pouvoir courir une chance sur 20.000 d'être en perte au bout de 9.000 coups contre un joueur qui parie pour la sortie d'un numéro, il se dira que le cas le plus défavorable est que ce joueur gagne 140 fois ; il peut donc lui promettre de lui payer pour chaque coup gagné $\frac{9000}{140}$ = environ 64 fois la mise ; le jeu serait *équitable* s'il payait 90 fois la mise.

Répétons encore une fois que la *règle de l'unité décimale* est une forme simplifiée et approchée d'une formule plus rigoureuse qui se déduit elle-même par des transformations analytiques, sans aucune hypothèse, des formules de dénombrement des diverses possibilités (au moyen de l'analyse combinatoire), dénombrement dont nous avons donné des exemples.

31. — On donne le nom de problèmes de *probabilités continues*, ou *probabilités géométriques*, aux problèmes dans lesquels le nombre des cas possibles est égal au nombre des positions possibles d'un point sur une droite, ou dans un plan, ou au nombre de positions d'une droite dans l'espace, etc. Etudions d'abord le cas le plus simple.

Considérons un segment de droite AB et un point M que nous supposons *assujetti à rester sur* AB (fig. 2). Quelle est la probabilité pour que le point M occupe sur AB une position déterminée ? On voit que *le nombre des cas possibles est indéfini*, car le point M peut varier d'une manière continue entre A et B ; il est donc nécessaire de donner une définition nouvelle de la probabilité. Une

telle définition est évidemment une *convention* pour le mathématicien, car celui-ci peut raisonner sur n'importe quelle définition ; mais ce n'est pas une *convention arbitraire;* elle est suggérée *a priori* par l'étude pratique de diverses questions et vérifiée *a posteriori* par la conformité de ses conséquences avec les observations. Quand le mathématicien dit que c'est une convention, cela veut dire simplement que, n'étant pas en état de *démontrer cette définition* d'une

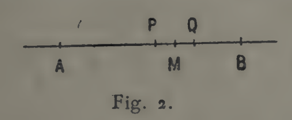


Fig. 2.

manière rigoureuse, il préfère la considérer comme une convention et rendre ainsi ses raisonnements abstraits absolument inattaquables. La difficulté qui consiste à légitimer la convention se trouve ainsi transportée au début de l'étude de chaque cas particulier concret.

La définition que nous adopterons est la suivante.

DÉFINITION. — *La probabilité pour que le point* M *se trouve sur un certain segment* PQ *de* AB *est proportionnelle à la longueur de ce segment.*

Conséquence. — Si on suppose M assujetti à se trouver sur AB, la probabilité pour que M soit sur AB est 1; donc *la probabilité pour que* M *soit sur* PQ *est égale au rapport* $\frac{PQ}{AB}$.

Nous discuterons tout à l'heure les objections et les critiques que l'on peut faire à cette définition, ainsi qu'aux définitions analogues pour les

probabilités géométriques dans le plan et dans l'espace ; notre conclusion sera que ces définitions sont légitimes, en ce sens qu'elles conduisent à des conséquences en accord avec l'expérience, du moment que les conditions de l'expérience ne sont pas compliquées exprès pour les mettre en défaut. Mais, sans attendre cette discussion critique, nous avons le droit d'étudier les conséquences mathématiques de cette définition de la probabilité continue ; nous arriverons ainsi à des conséquences dont la valeur sera celle même de la définition dont nous serons partis.

La plupart des problèmes de probabilités continues sont du ressort du calcul intégral ; pour cette raison, je me contenterai d'indiquer les résultats en omettant les démonstrations ; voici cependant quelques exemples où l'on peut aboutir en ne faisant usage que de procédés élémentaires.

32. — *Deux points M et M′ sont assujettis à se trouver sur un segment AB, de longueur l; quelle est la probabilité pour que la distance MM′ soit inférieure à kl, k étant donné inférieur à un.*

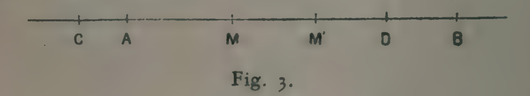
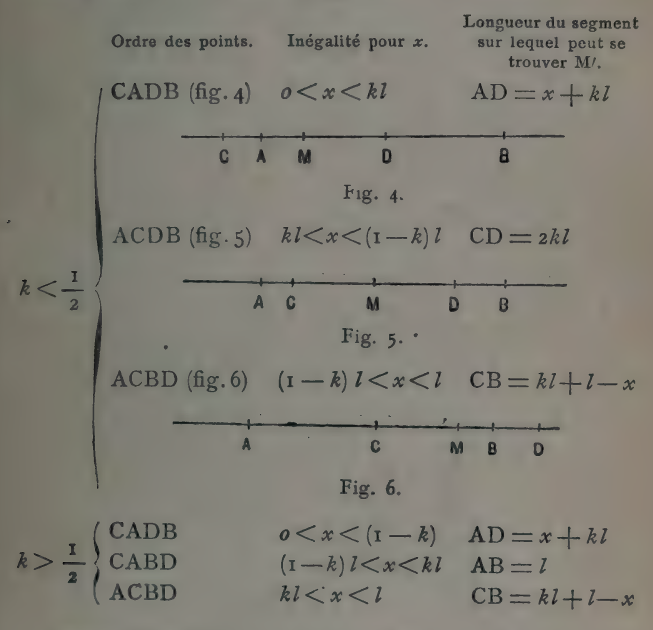


Fig. 3.

Désignons par x la distance AM ; la probabilité pour que cette distance soit comprise entre x et $x + dx$ est $\frac{dx}{l}$. Prenons (fig. 3) de part et d'autre du point M deux longueurs, $MC = MD = kl$; pour que MM′ soit inférieur à kl il faut et

il suffit que M′ soit compris entre C et D ; mais M′ doit aussi être compris entre A et B ; il faut donc distinguer plusieurs cas suivant les positions relatives des points C, A, D, B ; comme le segment CD est égal à $2kl$, il sera inférieur ou supérieur à AB suivant que k sera inférieur ou supérieur à $\frac{1}{2}$; nous examinerons successivement ces deux cas, à chacun desquels correspondent 3 cas différents de figure.



	Ordre des points.	Inégalité pour x.	Longueur du segment sur lequel peut se trouver M′.
$k < \frac{1}{2}$	CADB (fig. 4)	$o < x < kl$	$AD = x + kl$
	ACDB (fig. 5)	$kl < x < (1-k)l$	$CD = 2kl$
	ACBD (fig. 6)	$(1-k)l < x < l$	$CB = kl + l - x$
$k > \frac{1}{2}$	CADB	$o < x < (1-k)$	$AD = x + kl$
	CABD	$(1-k)l < x < kl$	$AB = l$
	ACBD	$kl < x < l$	$CB = kl + l - x$

Fig. 4.

Fig. 5.

Fig. 6.

Représentons le segment $AB = l$ (fig. 7 et 8) et élevons en chaque point M de AB une ordonnée MQ égale à la longueur du segment que peut occuper M′ lorsque M est au pied de l'ordonnée, le quotient par l^2 de l'aire (couverte de hachures)

ainsi obtenue représente la probabilité cherchée.

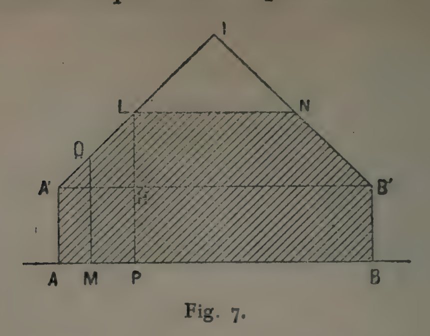


Fig. 7.

Dans la figure 7, correspondant à $k < \frac{1}{2}$, on a

$AA' = lk$, $AH = HL = kl$ $PL = 2kl$ $LN = l(1 - 2k)$

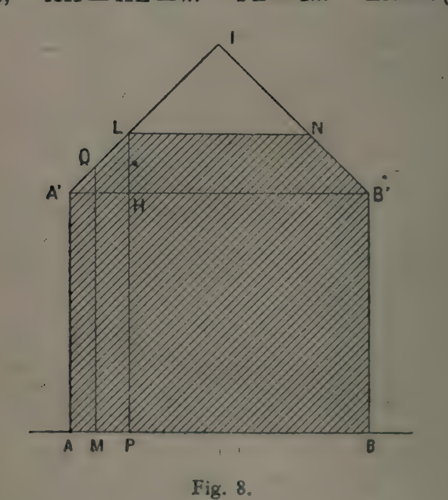


Fig. 8.

Dans la figure 8, correspondant à $k > \frac{1}{2}$, on a

$AA' = kl$ $A'H = HL = (1 - k)l$ $IL = l$ $LN = l(2k - 1)$

Sur les deux figures, l'aire couverte de hachures est égale à rectangle AA′B′B + triangle A′B′I — triangle LNI, c'est-à-dire à :

$$AB \times AA' + \frac{1}{4}\overline{A'B'}^2 - \frac{1}{4}\overline{LN}^2.$$

Mais, dans les deux figures, on a

$$AB \times AA' = l^2 k$$
$$\overline{A'B'}^2 = l^2$$
$$\overline{LN}^2 = l^2 (1 - 2k)^2$$

car LN est égal à $\pm l (1 - 2k)$; la probabilité qui s'obtient en divisant l'aire par l^2, est donc

$$k + \frac{1 - (1 - 2k)^2}{4} = 2k - k^2.$$

Indiquons une autre méthode pour arriver au même résultat ; posons

$$AM = x \qquad AM' = y.$$

La probabilité pour que AM soit compris entre $x + dx$ et M′ entre $y + dy$ est égale à $\frac{dxdy}{l^2}$. Si donc on représente le point P dont les coordonnées rectangulaires sont x et y (fig. 9) le point sera intérieur à un carré OABC de côté l, et la probabilité pour qu'il soit dans une aire donnée est proportionnelle à cette aire. Or la condition imposée par l'énoncé est que l'on ait

Fig. 9.

$$|x - y| < kl$$

ce qui exige que P soit compris entre les droites DE et FG, parallèles à OB et ayant pour équations respectives

$$x - y = \pm kl.$$

La probabilité cherchée est donc égale au rapport de l'aire ODEBGFO à l'aire du carré ; or les deux triangles rectangles ADE et CFG forment, par leur réunion, un carré de côté $(1 - k)\, l$ (car on a $OD = OF = kl$) ; la probabilité cherchée est donc

$$\frac{l^2 - (1 - k)^2 l^2}{l^2} = 2k - k^2.$$

Comme il fallait s'y attendre la probabilité croît constamment lorsque k varie de 0 à 1 ; pour $k = 0$ elle est égale à 0 et pour $k = 1$ elle est égale à 1 ; pour $k = \frac{1}{2}$, elle est égale à $\frac{3}{4}$.

33. — L'un des problèmes les plus célèbres de probabilités continues est le *problème de l'aiguille* que Buffon a étudié théoriquement et expérimentalement et dont il a le premier donné une solution correcte.

Sur un plan horizontal sont tracées des parallèles équidistantes ; on y lance au hasard une aiguille cylindrique ; quelle est la probabilité pour que l'aiguille rencontre l'une des parallèles ?

Désignons par $2a$ la distance des parallèles et par $2l$ la longueur de l'aiguille ; on suppose généralement l inférieur à a, de sorte qu'il ne peut y avoir qu'un point d'intersection. Nous nous bornerons à ce cas.

Un raisonnement ingénieux, dû à Barbier[1],

1. *Journal de Liouville*, 1860 (2e série, t. V)

permet de calculer sans peine l'espérance mathématique du joueur qui doit recevoir 1 franc pour chaque point d'intersection de l'aiguille avec les parallèles ; dans le cas $l < a$, comme il ne peut y avoir plus d'un de ces points, l'espérance mathématique se confond avec la probabilité.

Considérons une ligne polygonale (ouverte ou fermée) supposée matérialisée et jetons-la au hasard sur les parallèles ; il est clair que l'espérance mathématique du joueur qui doit recevoir 1 franc pour chaque point d'intersection est égale à la somme des espérances mathématiques pour les divers côtés du polygone, lesquelles sont égales elles-mêmes à la somme des espérances analogues pour les segments égaux en lesquels on peut diviser les côtés (supposés commensurables par le mode de raisonnement classique). On en conclut que *l'espérance mathématique cherchée est proportionnelle à la longueur du polygone*. Or, si celui-ci devient un cercle de diamètre $2a$, elle est évidemment égale à 2, car il y a toujours 2 points d'intersection et 2 seulement ; la longueur d'un tel cercle est $2\pi a$; l'espérance mathématique *par unité de longueur* est donc

$$\frac{2}{2\pi a} = \frac{1}{\pi a}.$$

Pour une aiguille de longueur $2l$ on retrouve le résultat $\frac{2l}{\pi a}$ donné par Buffon ; dans le cas où $a = 2l$, c'est-à-dire où la distance des parallèles est le double de la longueur de l'aiguille, la probabilité est l'inverse du rapport de la circonférence au diamètre.

Le problème de l'aiguille est le type d'un pro-

blème de probabilités continues dans lequel la « vérification expérimentale » est aisée ; en prenant une aiguille à coudre de dimension moyenne et la jetant au hasard sur une feuille de papier à dessin sur laquelle sont tracées des parallèles équidistantes (la distance étant, pour fixer les idées, le double de la longueur de l'aiguille), on trouve généralement, en quelques centaines d'épreuves, une valeur du nombre π assez approchée (en prenant le quotient du nombre total des épreuves par celui des épreuves où l'aiguille coupe l'une des parallèles). Il faut avoir soin, bien entendu, d'éviter de jeter l'aiguille de telle manière que sa direction soit, ou sensiblement parallèle, ou sensiblement perpendiculaire, à la direction des parallèles tracées sur le papier ; une bonne méthode consiste à lancer l'aiguille en l'air de telle manière qu'elle retombe à peu près verticalement sur une de ses extrémités ; la direction qu'elle prend est ainsi influencée le moins possible par la manière particulière adoptée pour la lancer.

34. — Joseph Bertrand, dans son célèbre ouvrage sur le *Calcul des Probabilités*, a critiqué la théorie des probabilités continues. Sa critique porte sur la définition même de la probabilité élémentaire, du moment qu'il y a une infinité de cas possibles. On sait qu'un nombre x est compris entre 0 et 10, quelle est la probabilité pour qu'il soit compris entre 0 et 5 ? Le bon sens conduit à répondre que cette probabilité est $\frac{5}{10}$ ou $\frac{1}{2}$ mais ce prétendu bon sens est illusoire, observe Bertrand, car si nous considérons le carré de x,

il est clair que si x est compris entre o et 10, son carré est compris entre o et 100 et si x est compris entre o et 5, son carré est compris entre o et 25 : le même bon sens nous donnera donc comme probabilité de cette dernière éventualité $\frac{25}{100}$ ou $\frac{1}{4}$. On trouverait, bien entendu, une autre valeur encore pour la probabilité en considérant le cube de x, ou le logarithme de x, ou une autre fonction de x. La conclusion de Bertrand paraît être que les problèmes de probabilités continues sont de purs jeux mathématiques, mais ne correspondent à aucune réalité; la définition de la probabilité élémentaire est arbitraire ; de cette définition arbitraire on tire des conséquences logiques, mais ces conséquences participeront toujours de l'arbitraire de la définition.

Cette attitude sceptique n'est pas en accord avec les faits. On peut s'en rendre compte aisément en empruntant un exemple même de Bertrand.

PROBLÈME. — *On trace une corde au hasard dans un cercle; quelle est la probabilité pour que sa longueur soit supérieure au côté du triangle équilatéral inscrit?*

Bertrand cite ce problème comme exemple d'un énoncé incomplet; il en donne en effet trois solutions qui conduisent à des résultats différents.

Première solution. — On peut, pour des raisons de symétrie, se donner la direction de la corde; le point d'intersection de cette corde avec le diamètre perpendiculaire à cette direction devra alors se trouver sur un segment égal à la moitié

de la longueur de ce diamètre (car la distance au centre du côté du triangle équilatéral inscrit est égale à la moitié du rayon) ; la probabilité est donc $\frac{1}{2}$.

Deuxième solution. — On peut, pour des raisons de symétrie, se donner une des extrémités de la corde sur le cercle ; la tangente en ce point et les deux côtés du triangle équilatéral inscrit ayant ce point pour sommet forment trois angles de 60° ; la direction de la corde doit être à l'intérieur de l'un de ces trois angles à l'exclusion des deux autres ; la probabilité est donc $\frac{1}{3}$.

Troisième solution. — Pour fixer la position d'une corde, il suffit de donner son milieu ; pour que la corde satisfasse à la condition de l'énoncé, il faut que son milieu soit intérieur à un cercle concentrique au cercle donné et de rayon moitié ; la surface de ce cercle est le quart de la surface du cercle donné ; la probabilité est donc $\frac{1}{4}$.

Doit-on penser que ces trois solutions sont également bonnes et, par suite, également mauvaises ? Nullement, il s'agit simplement de préciser le mode d'après lequel se fera la vérification expérimentale, c'est-à-dire comment on s'y prendra pour tracer une corde *au hasard* dans un cercle : si l'on assujettit cette corde à passer par un point fixe du cercle, ou si l'on fixe son milieu *au hasard*, c'est la seconde ou la troisième solution qui est la bonne ; mais il est aisé de voir que la plupart des procédés naturels que l'on peut imaginer conduisent à la première.

Si, par exemple, on jette un disque circulaire sur un plan sur lequel sont tracées des droites, la

probabilité pour que l'une des cordes interceptées soit supérieure au côté du triangle équilatéral est égale à $\frac{1}{2}$; il en est de même si l'on considère les cordes interceptées sur le disque circulaire de la lune par la trajectoire d'une étoile occultée, ou les cordes décrites dans le champ circulaire d'une lunette astronomique par des astres que l'on n'a pas visés, et qui occupent par suite une position quelconque dans le champ de la lunette.

On peut aisément[1] traiter un problème qui est la généralisation du précédent (probabilité pour qu'une corde d'un cercle ait une longueur supérieure à une longueur donnée) et déduire de la solution de ce problème celle du problème de l'aiguille. On obtiendra ainsi pour le problème de l'aiguille trois solutions différentes suivant que l'on aura choisi la première, la seconde ou la troisième des solutions proposées par Bertrand. Comme la solution expérimentale du problème de l'aiguille n'est pas contestée on est conduit, dans ce cas pratique, à constater que la première des solutions de Bertrand est la seule bonne ; il ne nous appartient donc pas, en modifiant arbitrairement une définition, de modifier la solution du problème.

35. — Voici un autre exemple de problème de probabilités continues, dans lequel nous verrons aussi que l'attitude de Bertrand est trop sceptique. Ce problème est un exemple des problèmes de probabilités relatives à la position des points sur une sphère.

1. Voir mes *Eléments de la théorie des probabilités*, p. 111.

Ces problèmes ont une certaine importance, car on peut y rattacher, comme nous le verrons au chapitre suivant, diverses questions relatives à la position des étoiles sur la sphère céleste.

PROBLÈME. — *Deux points* M *et* M′ *sont pris au hasard sur la surface de la sphère; quelle est la probabilité pour que le plus petit arc de grand cercle* MM′ *soit inférieur à* α?

La probabilité sera la même quelle que soit la position de M; or lorsque le point M est fixé, M′ doit se trouver sur une calotte sphérique AB, correspondant à un demi-angle au centre MOA $= \alpha$ (fig. 10). On a donc, en désignant par R le rayon de la sphère

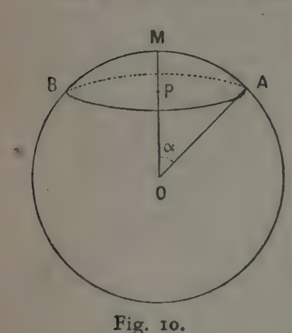


Fig. 10.

$$\begin{aligned} MP &= OM - OP \\ &= R\,(1 - \cos\alpha). \end{aligned}$$

Le rapport de l'aire de la calotte à l'aire de la sphère est

$$\frac{MP}{2R} = \frac{1 - \cos\alpha}{2} = \sin^2\frac{\alpha}{2}.$$

Telle est la probabilité cherchée; si α est très petit, on peut remplacer le sinus par l'arc et prendre pour valeur approchée $\frac{\alpha^2}{4}$.

Bertrand indique en même temps que la méthode précédente un autre raisonnement qui conduit à un résultat tout différent[1]. Lorsqu'on

1. J. BERTRAND. Calcul des probabilités. Chapitre I.

donne deux points M et M′ l'arc de grand cercle qui les joint est déterminé[1] ; tous les arcs de grand cercle étant analogues sur la sphère, on ne change pas la probabilité en fixant cet arc de grand cercle : or la probabilité pour que deux points M et M′ d'un cercle soient tels que l'arc MM′ soit inférieur à α est, comme nous l'avons vu $\frac{\alpha}{\pi}$; ce résultat est très différent du précédent, surtout si α est très petit ; si la mesure de α est 1°, on a $\alpha = \frac{\pi}{180}$ et par suite :

$$\frac{\alpha^2}{4} = \frac{\pi^2}{360^2} \qquad \frac{\alpha}{\pi} = \frac{1}{180}.$$

Le rapport de la seconde valeur à la première est $\frac{720}{\pi}$, c'est-à-dire plus de 200. Doit-on en conclure avec Bertrand que le problème proposé ne peut pas être résolu et que la première solution que nous en avons donnée est incorrecte ? Cette solution est, au contraire, la seule correcte, si l'on admet le postulat relatif à la probabilité élémentaire, c'est-à-dire si l'on considère toutes les portions égales de la sphère comme équivalentes entre elles, au point de vue de la probabilité qu'a le point M ou le point M′ de s'y trouver. Or, dans son second raisonnement, Bertrand paraît bien admettre cette sorte d'*homogénéité*, si l'on peut ainsi dire, qui fait regarder tous les points de la sphère comme équivalents entre eux. Ce second raisonnement renferme donc une cause d'inexactitude, qu'il n'est pas inutile de mettre en évidence. Examinons-le d'un peu près.

1. On peut écarter le cas dont la probabilité est évidemment nulle, où M et M′ sont diamétralement opposés

On commence par affirmer que la probabilité n'est pas changée lorsqu'on fixe l'arc de grand cercle sur lequel se trouvent M et M′ et la position de M sur ce grand cercle : cela est évident par raison de symétrie. L'erreur commence lorsque, le point M et le grand cercle étant fixés, on considère que la probabilité pour que M′ se trouve sur un arc donné de ce grand cercle est proportionnelle à la longueur de cet arc. Si l'arc de grand cercle est sans épaisseur il faut, pour parler rigoureusement, assigner la valeur zéro à la probabilité pour que M et M′ soient sur ce cercle ; on doit, pour ne pas avoir ce facteur zéro qui rend tout calcul impossible, considérer un faisceau délié d'arcs de grand cercle passant par un même point M et il est alors visible que la probabilité est plus grande pour M′ situé à un quadrant de M que situé au voisinage de M (fig. 11). Supposons, par exemple, que la sphère soit la sphère terrestre, M le pôle nord et M′ le point de chute d'un bolide qui se meut dans l'espace suivant une loi inconnue, avec une vitesse beaucoup plus grande que la vitesse de la terre. On demande quelle est la probabilité pour que le point de chute de M′ ait une latitude nord supérieure à 89° ; le raisonnement de Bertrand revient à dire : on peut supposer connue la longitude du point de chute M′, supposer, par exemple, que M′ se trouve sur le méridien de Paris ; tous les points de ce méridien sont alors également pro-

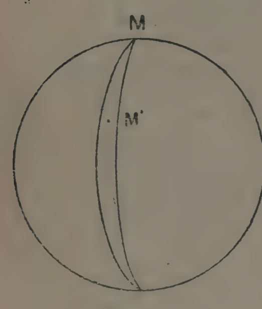


Fig. 11.

bables. Nous allons voir que, même en acceptant ce point de vue, on arrivera, en serrant de près la vérification expérimentale, à constater que tous les points du méridien ne sont pas également probables : comment, en effet, constatera-t-on qu'un point M′ est sur le méridien de Paris ? En déterminant sa longitude à l'aide d'observations astronomiques et d'un chronomètre ; cette détermination se fait avec une certaine précision angulaire ; supposons que cette précision soit de 0″,1, par exemple : cela veut dire que l'on considère, non pas la ligne idéale sans épaisseur que serait le méridien théorique, mais l'espace compris entre 0″,1 de longitude est et 0″,1 de longitude ouest, et cet espace est beaucoup plus grand à l'équateur, qu'au voisinage du pôle : la probabilité pour que la latitude de M′ soit comprise entre 0° et 1° est donc beaucoup plus grande que la probabilité pour que cette latitude soit comprise entre 89° et 90°.

On pourrait penser que, pour réfuter le raisonnement précédent, il suffirait de substituer aux observations astronomiques pour la détermination de la longitude, des mesures géodésiques. Supposons donc que, par des mesures d'une grande précision, on soit arrivé à tracer un méridien sur la surface terrestre ; on aura du moins fixé un certain nombre de repères, suffisamment rapprochés pour que deux au moins soient aperçus de tout point du méridien et, sur chacun de ces repères un trait vertical extrêmement délié fait connaître son intersection avec le plan du méridien. On supposera alors que le centre de gravité M′ du bolide se trouve sur le méridien ainsi défini ; la probabilité ne sera-t-elle pas la même à l'équateur qu'au pôle ? Oui, sans doute,

si les traits verticaux tracés sur les repères ont partout la même épaisseur, mais cette épaisseur uniforme des traits est contradictoire avec la notion du méridien : quelle que soit l'épaisseur choisie pour le trait du repère situé à l'équateur (disons, par exemple, un dixième de millimètre), les deux bords de ce trait définissent avec le pôle *deux méridiens*, dont l'angle est sans doute très petit, puisque leur écartement à l'équateur est seulement un dixième de millimètre, mais cet écartement est encore plus petit lorsqu'on s'approche du pôle; si l'on plaçait à l'équateur deux traits verticaux voisins, ayant chacun un dixième de millimètre d'épaisseur, le bord droit de l'un coïncidant avec le bord gauche de l'autre, il ne serait pas possible de prolonger jusqu'au pôle les méridiens ainsi définis, *en leur conservant leur largeur de un dixième de millimètre :* si on parvenait à réaliser cette construction géodésique infiniment délicate, les traits d'un dixième de millimètre qui seraient tracés à une latitude quelconque *empiéteraient l'un sur l'autre*, d'autant plus que la latitude s'approcherait davantage de 90°, de sorte que, dans la région commune à ces traits, on n'aurait aucune raison pour dire que l'on se trouve sur l'un ou sur l'autre des méridiens définis à l'équateur.

Cette discussion un peu longue n'était pas inutile pour réduire à sa juste valeur la boutade de Bertrand, à laquelle il n'attachait pas lui-même plus de valeur qu'il ne convient [1]. On doit la con-

1. Voir Darboux, Eloge historique de Joseph Bertrand. A propos du paradoxe analogue, dont nous avons parlé tout à l'heure, M. Darboux dit que Bertrand connaissait la solution, mais préférait la laisser deviner à son lecteur.

sidérer comme un exemple mettant en garde contre des raisonnements inexacts.

36. — Henri Poincaré a repris la discussion des paradoxes de Joseph Bertrand, avec l'esprit de généralisation analytique et la forme particulière de critique qui constituent le fond de sa philosophie.

Tout d'abord, Poincaré observe que l'on peut à une variable continue x substituer une fonction continue *quelconque* $f(x)$ de cette variable et remplacer le problème de probabilité relatif à x par le problème correspondant relatif à $f(x)$; la même remarque s'étend au cas où l'on a plusieurs variables au lieu d'une seule; on en tire immédiatement la conclusion que la définition de la probabilité élémentaire est *entièrement arbitraire*, puisque l'on peut faire varier cette définition d'après une loi qui dépend d'une fonction continue *absolument quelconque*[1]. Arrivé là,

1. Si la position d'un point dépend de deux variables, on introduira une fonction arbitraire $\varphi(x, y)$ assujettie à la double condition d'être positive et de vérifier la relation

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \varphi(x, y)\, dx\, dy = 1$$

et la probabilité pour que le point x, y soit dans une aire (S) s'exprimera par l'intégrale double

$$\iint_{(S)} \varphi(x, y)\, dx\, dy.$$

D'une manière générale si l'on pose :

$$x = f(\alpha, \beta)$$
$$y = g(\alpha, \beta)$$

et

$$\Phi(\alpha, \beta) = \varphi(f, g)\frac{D(f, g)}{D(\alpha, \beta)}$$

tout autre que Poincaré se serait arrêté, semble-t-il, puisque le résultat est entièrement négatif. C'est ici au contraire, que Poincaré introduit une idée neuve et très intéressante.

37. — Quel que soit, en effet, l'intérêt analytique des calculs sur les fonctions arbitraires, il ne saurait suffire à justifier l'introduction de ces fonctions arbitraires s'il ne se produisait fréquemment une circonstance très remarquable, sur laquelle Poincaré a le premier attiré l'attention : dans certains cas, le résultat final du calcul est, dans une large mesure, *indépendant du choix de la fonction arbitraire*; il suffit de supposer que cette fonction satisfait à des conditions assez lar-

la probabilité s'exprimera par la formule

$$\iint_{\Sigma} \Phi(\alpha, \beta)\, d\alpha\, d\beta$$

en désignant par Σ l'aire (supposée ne pas se recouvrir elle-même) qui correspond dans le champ des variables α, β à l'aire S dans le champ des variables x, y.

On peut, d'une infinité de manières (sous des conditions de continuité pour φ que nous n'énoncerons pas), déterminer le changement de variables de manière que l'on ait

$$\Phi(\alpha, \beta) = 1.$$

La probabilité devient alors

$$\int_{\Sigma} d\alpha\, d\beta$$

c'est-à-dire est proportionnelle à l'aire Σ. Ces variables α, β ainsi déterminées seront dites variables normales pour la fonction $\varphi(x, y)$ qu'on s'est donnée *a priori*. Ainsi, pour toute fonction arbitraire, il y a des variables convenablement choisies qui rendent cette fonction arbitraire égale à l'unité et qui, par suite, peuvent justifier par leur emploi le choix en apparence arbitraire. Mais il ne faut pas perdre de vue que l'arbitraire dans le choix des variables n'a pas plus de base concrète que l'arbitraire dans le choix de la fonction $\varphi(x, y)$; à un problème déterminé correspondent des variables bien déterminées, qui ne dépendent que de la manière dont est posé le problème.

ges relatives à sa continuité ou au sens de sa variation. Nous allons nous rendre compte de ce fait important par l'étude d'un problème simple.

PROBLÈME. — *Un cadran circulaire est divisé en 2n parties égales, peintes alternativement en blanc ou en noir ; une aiguille est mobile autour du centre du cadran et est lancée avec une force suffisante pour faire, avant de s'arrêter, plusieurs fois le tour du cadran ; quelle est la probabilité pour qu'elle s'arrête en face d'une division blanche du cadran* ?

On remarquera que l'énoncé ne mentionne pas la position initiale de l'aiguille ; le résultat est en effet, sensiblement indépendant de cette position grâce à l'hypothèse sur la force avec laquelle est lancée l'aiguille ; si celle-ci avait un élan très faible, ne lui faisant parcourir que deux ou trois divisions, il n'en serait pas de même.

Désignons par θ l'angle total dont tourne l'aiguille ; cet angle augmente de 2π à chaque tour complet du cadran ; d'après l'hypothèse, il dépasse toujours les premiers multiples entiers de 2π. Supposons que ce soit toujours une même personne qui lance l'aiguille : nous pourrons désigner par $\varphi(\theta)\, d\theta$ la probabilité pour que l'angle total dont elle tourne soit compris entre θ et $\theta + d\theta$; la fonction $\varphi(\theta)$ est nulle lorsque θ est inférieur à 2π, puisque l'aiguille fait toujours plus d'un tour ; elle sera nulle aussi lorsque θ dépassera $100\,\pi$, si l'on suppose que la personne qui lance l'aiguille, en employant toute sa force, est incapable de faire faire à l'aiguille plus de 50 tours ; lorsque θ sera compris entre 2π et 100π,

nous ne ferons sur la fonction $\varphi(\theta)$ que des hypothèses de *continuité*; ces hypothèses sont naturelles si l'on suppose l'appareil bien construit au point de vue de l'égalité des résistances passives qui ralentissent le mouvement de l'aiguille ; si, par suite d'une défectuosité de l'appareil, une résistance plus grande se produisait dans le voisinage d'une certaine valeur de θ, les arrêts

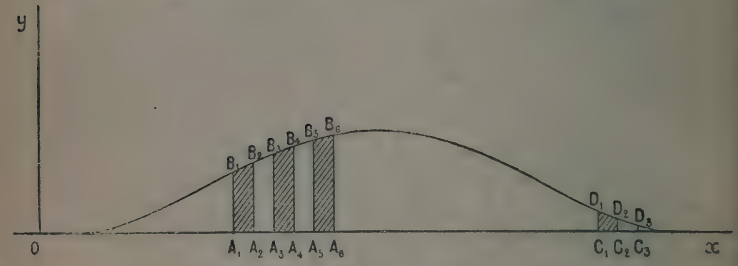


Fig. 12.

seraient plus fréquents dans cette région et, si cette région correspondait à une division blanche du cadran, la probabilité pour l'arrêt en face de cette division se trouverait accrue.

Notre hypothèse sur θ revient donc à ceci ; la fonction

$$y = \varphi(x)$$

peut être représentée par une courbe continue telle que celle que nous figurons (fig. 12) ; on a d'ailleurs

$$\int_{-\infty}^{+\infty} \varphi(x)\, dx = 1.$$

Nous avons figuré cette courbe avec un seul maximum ; le raisonnement serait analogue s'il y en avait plusieurs, en nombre limité. Dans les

deux cas, on arrive aisément à prouver, en se basant seulement sur la continuité, que le rapport de la somme des aires couvertes de hachures, dans la figure 12, à la somme des aires non hachurées, diffère très peu de l'unité, du moment que les ordonnées sont suffisamment nombreuses. La probabilité demandée est donc égale à $\frac{1}{2}$.

Ce résultat est fort intéressant, mais son intérêt ne doit pas masquer ce qu'il y a de conventionnel dans une introduction trop systématique des fonctions arbitraires.

Du moment que ces problèmes ont un sens concret, c'est-à-dire sont susceptibles d'une étude expérimentale, la définition à adopter pour la probabilité élémentaire, est une conséquence nécessaire de la technique expérimentale adoptée ; il n'y a donc pas place pour une définition arbitraire. C'est dans certaines questions de probabilités des causes que l'artifice analytique de Poincaré semble rendre le plus de services.

L'attitude de Poincaré, on s'en est rendu compte, n'est sceptique qu'en apparence ; si la définition de la probabilité continue ne lui paraît pas absolument satisfaisante, c'est parce qu'il préfère toujours une certitude abstraite à une certitude concrète et, lorsque l'introduction de la fonction continue arbitraire lui a permis d'atteindre ce degré supérieur de certitude, il est satisfait de ce progrès et ne s'inquiète pas de savoir si l'excès de ses exigences ne risque pas de conduire au scepticisme des esprits moins puissants que le sien, alors qu'il ne les a témoignées que pour se donner la joie intellectuelle de les satisfaire.

CHAPITRE IV

PROBABILITÉ DES CAUSES

38. — On a donné le nom de problèmes de *probabilités des causes* à une catégorie de problèmes dans lesquels on se propose, connaissant un événement, de déterminer les probabilités d'événements antérieurs inconnus, mais en relation avec l'événement connu.

Ces événements antérieurs ne sont pas d'ailleurs supposés entièrement inconnus ; on suppose que l'on connaît ce que l'on appelle leur *probabilité a priori*, c'est-à-dire la probabilité qu'ils auraient si l'événement connu ne s'était pas produit ; on se propose de déterminer *leur probabilité a posteriori*, c'est-à-dire après cet événement.

Donnons quelques exemples simples.

Problème I. — *Deux urnes* A *et* B *renferment, l'urne* A: 3 *boules blanches et* 1 *noire ; l'urne* B, 1 *boule blanche et* 3 *noires. On extrait une boule d'une des urnes, choisie au hasard et on contate qu'elle est blanche ;*

quelle est la probabilité pour que l'urne choisie soit A?

Dans cet énoncé, l'indication d'après laquelle l'urne est *choisie au hasard,* signifie que la probabilité *a priori* est la même pour l'urne A et l'urne B ; elle est donc $\frac{1}{2}$ pour chacune d'elles.

Chacune des urnes renfermant le même nombre de boules, la probabilité d'extraire une quelconque des boules est la même ; comme il y a 3 boules blanches dans A et 1 dans B, la probabilité pour que la boule blanche extraite soit l'une des boules de A est $\frac{3}{4}$; telle est la probabilité demandée.

On peut aussi raisonner comme il suit : imaginons 40 systèmes de deux urnes A_1, A_2, ... A_{40} ; B_1, B_2, ... B_{40} couplées deux à deux, toutes les urnes A_i ont la même composition que l'urne A et toutes les urnes B_i ont la même composition que l'urne B. Supposons que l'on extraie une boule au hasard d'une urne de chaque couple choisie au hasard ; on peut présumer que l'urne choisie sera A dans 20 des couples et B dans les 20 autres ; or, si l'on extraie 20 fois une boule d'une urne dont la composition est A, l'hypothèse la plus probable est que l'on obtiendra 15 blanches et 5 noires ; de même on obtiendra 5 blanches et 15 noires dans les 20 extractions d'urnes B ; ce qui permet de former le tableau suivant

20 extr. A	donnent	15 blanches	5 noires
20 » B	»	5 »	15 »
40 »	»	20 »	20 »

On voit que les 20 extractions ayant fourni des blanches se décomposent en 15 provenant d'urnes

A et 5 provenant d'urnes B; la probabilité pour que l'une d'elles provienne d'une urne A est donc $\frac{15}{20}$ ou $\frac{3}{4}$.

Problème II. — *On a trois coffrets dont chacun a deux tiroirs; le premier coffret renferme dans chaque tiroir une pièce d'or, le second renferme une pièce d'argent et une pièce d'or et le troisième deux pièces d'argent. On ouvre l'un des tiroirs et on y trouve une pièce d'or; quelle est la probabilité pour que le second tiroir du même coffret renferme une pièce d'or?*

La question revient à la suivante; quelle est la probabilité pour que le tiroir ouvert appartienne au premier coffret? Or, trois tiroirs renfermaient une pièce d'or; la probabilité pour que l'on ait ouvert chacun d'eux est donc $\frac{1}{3}$, car ces 3 tiroirs sont supposés extérieurement identiques; la probabilité demandée est donc $\frac{2}{3}$ puisqu'il y a deux tiroirs à pièce d'or dans le premier coffret.

Ploblème III. — *On a dix urnes dont l'une* A *renferme* 5 *boules blanches et* 1 *noire, les* 9 *autres* B *renfermant chacune* 2 *blanches et* 2 *noires. On extrait une boule d'une urne choisie au hasard et on constate qu'elle est blanche. Quelle est la probabilité pour que l'urne choisie soit* A*?*

On ne peut pas ici raisonner aussi simplement que dans les problèmes précédents parce que les urnes B ne renferment pas le même nombre de

boules que l'urne A ; on peut, par l'emploi du plus petit multiple commun, être ramené à ce cas. Rien ne serait changé en effet, si l'on doublait le nombre des boules de l'urne A, en conservant leurs proportions et si l'on triplait de même le nombre des boules des urnes B ; l'urne A renferme alors 10 blanches et 2 noires et chaque urne B renferme 6 blanches et 6 noires ; il y a en tout $10 + 6 \times 9 = 64$ boules blanches parmi lesquelles 10 sont dans A ; la probabilité pour qu'une boule blanche étant extraite, l'urne A ait été choisie est donc $\frac{10}{64} = \frac{5}{32}$.

39. — La solution générale des problèmes de probabilités des causes analogues aux précédents est donnée par la formule de Bayes, que nous allons établir.

Considérons n urnes parmi lesquelles n_1 que nous désignerons indifféremment par A_1 ont la même composition et renferment une proportion p_1 de boules blanches ; de même n_2 urnes A_2 renferment une proportion p_2 de boules blanches, ..., n_k urnes A_k renferment une proportion p_k de boules blanches. Les $n = n_1 + n_2 + \dots + n_k$ urnes étant supposées identiques extérieurement, on choisit l'une d'elles et on en extrait une boule, qui se trouve être blanche. Quelle est la probabilité P_1 pour que l'urne choisie soit l'une des urnes A_1 ?

Nous supposerons que chaque urne renferme le même nombre N de boules ; les urnes A_1 renfermeront chacune p_1N boules blanches[1], les

1. Le nombre N est choisi tel que p_1 N, p_2 N soient des nombres entiers.

urnes A_2 en renfermeront chacune p_2N, etc., de sorte que le nombre total de boules blanches est

$$n_1p_1N + n_2p_2N + \dots + n_kp_kN$$

parmi lesquelles n_1p_1N appartiennent à l'une des urnes A_1. La probabilité P_1 pour que la boule blanche extraite provienne d'une urne A_1 est donc :

$$P_1 = \frac{n_1p_1N}{n_1p_1N + \dots + n_kp_kN} = \frac{n_1p_1}{n_1p_1 + n_2p_2 + \dots + n_kp_k}$$

On écrit souvent cette formule sous une autre forme en posant :

$$n_1 = \varpi_1 n \quad , \quad n_2 = \varpi_2 n \quad , \quad \dots . \, n_k = \varpi_k n.$$

La formule devient alors

$$P_1 = \frac{p_1\varpi_1}{p_1\varpi_1 + p_2\varpi_2 + \dots p_n\varpi_n}.$$

La signification des nombres ϖ_k est très simple. Si l'on choisit une urne au hasard, la probabilité pour que l'urne choisie soit de l'espèce A_1 est ϖ_1, puisque ϖ_1 est le rapport du nombre n_1 des urnes A_1 au nombre total n des urnes. On dira pour cette raison, que ϖ_1 est la *probabilité a priori* (c'est-à-dire avant tout tirage) pour que l'urne choisie soit de l'espèce A_1 ; la probabilité que nous avons calculée P_1 est alors dite la probabilité *a posteriori* et la formule précédente est la *formule de Bayes*. On peut la démontrer par le raisonnement suivant, où l'on fait usage des théorèmes sur les probabilités totales et composées.

Evaluons la probabilité pour que l'extraction donne *une boule blanche extraite de* A_1 ; nous évaluerons cette probabilité de deux manières en

intervertissant l'ordre des deux conditions auxquelles doit satisfaire la boule: 1° elle doit être blanche; 2° elle doit être extraite de A_1.

1° La probabilité pour que la boule extraite soit blanche est une probabilité totale, car la boule blanche peut être extraite, soit des urnes A_1, soit des urnes A_2, etc.; sa valeur est donc:

$$p_1\varpi_1 + p_2\varpi_2 + \cdots + p_n\varpi_n.$$

D'autre part, la probabilité pour qu'une boule extraite, *lorsqu'on sait qu'elle est blanche*, provienne de A_1 a été désignée par P_1; la probabilité demandée est donc

$$P_1(p_1\varpi_1 + p_2\varpi_2 + \cdots + p_n\varpi_n). \tag{1}$$

2° La probabilité que la boule extraite provienne de A_1 est égale à ϖ_1; sachant qu'elle vient de A_1, la probabilité pour qu'elle soit blanche est p_1; la probabilité demandée est donc

$$p_1\varpi_1. \tag{2}$$

En égalant les deux valeurs trouvées (1) et (2), on retrouve bien la formule de Bayes. Cette dernière démonstration s'applique quelle que soit la nature des expériences dans lesquelles les probabilités sont mises en jeu; ce qui rend parfois les applications difficiles, c'est l'incertitude qui règne sur la valeur des probabilités *a priori*. Donnons-en quelques exemples.

PROBLÈME IV. — *Une urne* A *renferme un nombre inconnu de boules blanches ou noires; on fait plusieurs tirages successifs, en remettant la boule extraite après chaque tirage; le résultat de ces tirages est l'extrac-*

tion de r *boules blanches et de* s *boules noires. Quelle est la composition la plus probable de l'urne*?

On est tenté de répondre que la composition la plus probable est telle que le rapport du nombre des boules blanches au nombre des boules noires est égal au rapport de r à s : mais ce résultat ne sera exact que si l'on fait sur les probabilités *a priori* des hypothèses non seulement très particulières, mais variables avec les valeurs de r et s ; c'est ce que montrera l'étude d'un exemple particulier.

Supposons que l'on ait $r = 2$, $s = 0$, c'est-à-dire que deux tirages consécutifs aient donné des boules blanches, et que l'on sache que l'urne renferme 6 boules. On peut faire sur sa composition 7 hypothèses que nous désignerons par A_0, A_1, ..., A_6, l'hypothèse A_k étant celle où l'urne renferme $6 - k$ blanches et k noires. Dans chacune de ces hypothèses la probabilité d'extraire deux blanches consécutives est donnée par le tableau suivant

A_0	6 blanches	0 noire	$p_0 = 1$
A_1	5 blanches	1 noire	$p_1 = \left(\frac{5}{6}\right)^2 = \frac{25}{36}$
A_2	4 blanches	2 noires	$p_2 = \left(\frac{4}{6}\right)^2 = \frac{16}{36}$
A_3	3 blanches	3 noires	$p_3 = \left(\frac{3}{6}\right)^2 = \frac{9}{36}$
A_4	2 blanches	4 noires	$p_4 = \left(\frac{2}{6}\right)^2 = \frac{4}{36}$
A_5	1 blanche	5 noires	$p_5 = \left(\frac{1}{6}\right)^2 = \frac{1}{36}$
A_6	0 blanche	6 noires	$p_6 = 0.$

Si donc on désigne par $\varpi_0, \varpi_1, \varpi_2, \ldots, \varpi_6$ les probabilités *a priori* de ces 7 hypothèses, leurs probabilités *a posteriori* $P_0, P_1, P_2, \ldots, P_6$ sont données par les formules

$$P_0 = \frac{36\varpi_0}{36\varpi_0 + 25\varpi_1 + 16\varpi_2 + 9\varpi_3 + 4\omega_4 + \varpi_5}$$

$$P_1 = \frac{25\varpi_1}{36\varpi_0 + 25\varpi_1 + 16\varpi_2 + 9\varpi_3 + 4\varpi_4 + \varpi_5}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$P_5 = \frac{\varpi_5}{36\varpi_0 + 25\varpi_1 + 16\varpi_2 + 9\varpi_3 + 4\varpi_4 + \varpi_5}$$

$$P_6 = 0.$$

On voit qu'il est possible de déterminer $\varpi_0, \varpi_1, \ldots, \varpi_5$ de manière que $P_0, P_1, \ldots, P_5$ aient des valeurs positives *quelconques* dont la somme soit égale à l'unité (on a forcément $P_6 = 0$, car si l'on a extrait des boules blanches, on n'est sûrement pas dans le cas où l'urne ne renferme que des noires).

Il est donc nécessaire de faire des hypothèses sur $\varpi_0, \varpi_1, \ldots, \varpi_5$, c'est-à-dire de compléter l'énoncé. Une hypothèse plausible consistera à supposer ces probabilités *a priori* égales entre elles, c'est-à-dire à admettre que les 7 compositions possibles sont représentées par 7 urnes entre lesquelles on choisit au hasard ; cette hypothèse donne :

$$P_0 = \frac{36}{91}, \quad P_1 = \frac{25}{91}, \quad P_2 = \frac{16}{91}, \quad P_3 = \frac{9}{91}, \quad P_4 = \frac{4}{91},$$
$$P_5 = \frac{1}{91}.$$

Une autre hypothèse consisterait à admettre que l'urne a été remplie au moyen de tirages effectués dans une urne renfermant en propor-

tions égales des boules blanches et noires; on a alors d'après le triangle arithmétique :

$$\varpi_0 = \varpi_6 = \frac{1}{64}$$

$$\varpi_1 = \varpi_5 = \frac{6}{64}$$

$$\varpi_2 = \varpi_4 = \frac{15}{64}$$

$$\varpi_3 = \frac{20}{64}$$

et par suite, on obtient

$$P_0 = \frac{6}{112}, \quad P_1 = \frac{25}{112}, \quad P_2 = \frac{40}{112}, \quad P_3 = \frac{30}{112},$$

$$P_4 = \frac{10}{112}, \quad P_5 = \frac{1}{112}.$$

On voit que ce résultat diffère beaucoup du précédent; l'hypothèse la plus probable est ici celle où l'urne renferme 4 blanches et 2 noires, parce que la probabilité *a priori* de cette hypothèse est bien plus grande que la probabilité *a priori* de l'hypothèse où l'urne renfermerait exclusivement des boules blanches.

40. — Il est des cas où l'évaluation de la probabilité *a priori* est une simple affaire d'appréciation ; le résultat est fonction de la valeur qui est attribuée à cette probabilité. En voici un exemple classique.

Pierre joue à l'écarté avec un inconnu qui, la première fois qu'il donne les cartes, retourne le roi. Quelle est la probabilité pour que le partenaire de Pierre soit tricheur de profession?

Dans la solution que Poincaré donne de ce problème, il admet implicitement qu'un tricheur de profession tourne le roi chaque fois qu'il en a l'occasion. Bien que je n'aie pas — plus que Poincaré — fréquenté « le monde où l'on triche », cette hypothèse me paraît assez peu vraisemblable; un joueur honnête tourne le roi en moyenne une fois sur huit; il paraît difficile qu'un individu à qui pareille bonne fortune arriverait à chaque coup n'attire pas rapidement l'attention; un tricheur, d'ailleurs, préférera varier ses moyens d'action, et, en tout cas, ne cherchera pas à forcer le sort lorsque celui-ci lui sera spontanément favorable. Il semble donc que l'on se rapproche de la vérité en évaluant à $\frac{1}{4}$, par exemple, la probabilité pour qu'un tricheur retourne le roi, alors que cette probabilité est $\frac{1}{8}$ pour le joueur honnête.

Désignons par ϖ la probabilité *a priori* pour que le partenaire de Pierre soit tricheur; la probabilité *a priori* pour qu'il soit honnête est $1 - \varpi$ et la probabilité P *a posteriori* (c'est-à-dire après qu'il a tourné le roi) pour qu'il soit tricheur est

$$P = \frac{\frac{1}{4}\varpi}{\frac{1}{4}\varpi + \frac{1}{8}(1 - \varpi)} = \frac{2\varpi}{1 + \varpi}.$$

Si $\varpi = \frac{1}{2}$, c'est-à-dire si l'on suppose égales les probabilités *a priori* pour que l'adversaire de Pierre soit tricheur et pour qu'il ne le soit pas, on trouve

$$P = \frac{2}{3}$$

c'est-à-dire que l'unique événement observé augmente notablement la probabilité. Mais l'on observe justement que Pierre serait bien imprudent d'engager une partie contre un partenaire de la moralité spéciale duquel il aurait une aussi fâcheuse opinion; le fait que Pierre joue suppose qu'il regarde ϖ comme un nombre assez petit. Si ϖ est petit, P diffère peu de 2ϖ; en tout cas d'ailleurs, P est inférieur à 2ϖ, c'est-à-dire que la probabilité est tout au plus doublée par l'événement observé; si le partenaire de Pierre est tel que ϖ soit égal à 1 sur 100.000, P est égal à 1 sur 50.000, c'est-à-dire pratiquement aussi négligeable que ϖ; si Pierre est sûr de son partenaire et suppose $\varpi = 0$, on a aussi $P = 0$, c'est-à-dire que l'événement ne doit pas ébranler sa confiance : les résultats du calcul sont conformes à ceux du bon sens.

41. — Les problèmes de probabilités des causes sont extrêmement variés et sont le plus souvent mélangés à d'autres problèmes de probabilités; on pourrait presque dire que dans tout problème réel de probabilités se trouve une question de probabilités des causes. C'est ce qui se produit notamment à propos des applications de la théorie des probabilités à la statistique et aux questions d'assurances.

Toutes les fois que la probabilité d'un événement simple est déterminée par l'expérience, la détermination de cette probabilité est en réalité un problème de probabilités des causes. On a observé, par exemple, que le nombre des naissances masculines est environ 51 p. 100 du nombre total des naissances; ce résultat statistique

est confirmé par les observations faites depuis un siècle environ dans les pays civilisés, c'est-à-dire par des statistiques qui portent sur environ un milliard de faits ; on peut donc se poser le problème suivant : un milliard d'extractions effectuées dans une urne de composition inconnue ayant fourni 510 millions de boules blanches et 490 millions de boules noires, quelle est la composition la plus probable de l'urne ? Un calcul que j'omets montre facilement qu'il est extrêmement peu probable que l'urne renferme autant de boules noires que de boules blanches, et ce résultat est, dans une large mesure, indépendant des hypothèses sur la probabilité *a priori* des diverses compositions possibles de l'urne.

Cet exemple montre la forme particulière que prennent souvent les problèmes de probabilités des causes. Cette forme est la suivante : *Tel résultat est-il dû au hasard, ou a-t-il une cause?* On a souvent fait observer combien cet énoncé manque de précision ; Bertrand a beaucoup insisté sur ce point : nous aurons l'occasion de citer les plus intéressantes de ses observations. Mais tout ce que l'on peut objecter au point de vue logique ne saurait empêcher la question précédente de se poser d'elle-même dans bien des cas : la théorie des probabilités ne peut donc refuser de l'examiner et d'y faire une réponse ; la précision de cette réponse sera naturellement limitée par le manque de précision de la question ; mais refuser de répondre sous prétexte que la réponse ne peut pas être absolument précise, c'est se placer sur un terrain purement abstrait et méconnaître le caractère qu'ont forcé-

ment les applications des mathématiques. Le calcul fournit, il est vrai, une réponse précise à toute question précise; mais, pratiquement, il n'y a jamais de question précise : les données expérimentales comportent forcément un certain jeu; la même imprécision affecte donc le résultat du calcul et la soi-disant précision théoriquement absolue de ce calcul n'est qu'une pure illusion.

Comme exemples de questions dans lesquelles la cause est inconnue, ou dans lesquelles *le problème est précisément de savoir s'il y a une cause*, on peut citer un grand nombre de recherches de psychologie expérimentale[1]; par exemple, on constate chez un certain nombre d'individus, la coïncidence entre une particularité physique et une particularité de nature psychologique; dans quelles limites peut-on en conclure à une corrélation réelle?

Signalons aussi les questions fort nombreuses que l'on peut se poser sur la langue d'un auteur; chaque écrivain a pour la longueur des mots et des phrases des habitudes particulières que l'on peut traduire en nombres; de telles remarques peuvent mettre sur la voie d'interpolations de textes. Dans d'autres cas, on pourra se poser et résoudre par la méthode statistique des questions de métrique[2].

On a aussi appliqué les méthodes de la théorie

1. Voir ch. VIII.

2. Les difficultés de ce genre de questions, et en même temps l'utilité réelle de la théorie des probabilités pour les résoudre ont été signalées avec beaucoup de justesse par M. Daniel Serruys dans une note trop technique pour que je puisse l'analyser ici : Les procédés toniques d'Himérius et les origines du « Cursus » byzantin (Mélanges Louis Havet : Philologie et linguistique, Paris, 1909).

des probabilités à des questions d'art : une école de sculpture, par exemple, est caractérisée par la fixité de certains rapports entre les mesures de diverses parties du corps ; des écarts numériques peuvent faire présumer une différence d'école que confirme une étude proprement esthétique[1].

Dans toutes ces questions, le calcul est surtout destiné à fournir des présomptions qu'il est nécessaire de mettre à l'épreuve d'une critique directe du problème posé, par la technique propre de son art.

42. — On peut se poser d'intéressants problèmes de probabilités des causes à propos de la distribution des étoiles sur la sphère céleste. Cette question est une des plus passionnantes de la philosophie naturelle, car elle touche aux origines et aux destinées de notre univers. Sauf quelques exceptions relativement rares, nous ignorons les distances des étoiles ; une première difficulté se présente donc : lorsque deux ou plusieurs étoiles sont rapprochées sur la carte du ciel, peut-on en conclure qu'elles sont rapprochées dans l'espace? Il ne saurait être question d'apporter à cette question une réponse rigoureuse ; mais l'étude des probabilités peut, dans certains cas, donner de très fortes présomptions en faveur de l'affirmative. A moins d'admettre, en effet, des alignements réguliers analogues à ceux de la théorie des réseaux cristallins, alignements dont aucune raison n'autorise à soupçonner l'existence, il est clair que, si deux ou plusieurs étoiles sont extrêmement éloignées dans l'espace, les probabilités

1. Voir Jean Laran : La méthode statistique dans un problème d'archéologie (*Revue du Mois* du 10 avril 1907, t. III, p. 433).

pour que leurs positions sur la sphère céleste soient rapprochées les unes des autres se calculeront d'après les principes du n° 45. Si dès lors on fait entrer en ligne de compte le nombre des étoiles de grandeur déterminée, par exemple le nombre des étoiles visibles à l'œil nu, ou avec telle lunette, on pourra calculer quelle est la probabilité pour qu'un nombre déterminé de ces étoiles soient situées à l'intérieur d'un petit cercle donné de la sphère céleste. S'il existe réellement un tel groupement dont la probabilité ainsi calculée soit très faible, on sera autorisé à penser que ce groupement a une *cause* autre que le hasard, c'est-à-dire qu'il est formé d'étoiles réellement rapprochées dans l'espace.

Bertrand fait à ce genre de raisonnement des objections qui ne me paraissent pas convaincantes :

« Les Pléiades, dit-il[1], semblent plus rappro-
« chées les unes des autres qu'il n'est *naturel*.
« L'assertion est digne d'intérêt; mais, si l'on
« veut traduire la conséquence en chiffres, les
« éléments font défaut. Faut-il, pour préciser
« cette idée vague de rapprochement, chercher
« le plus petit cercle qui contienne le groupe? la
« plus grande des distances angulaires? la somme
« des carrés de toutes les distances? l'aire du
« polygone sphérique dont quelques-unes des
« étoiles sont les sommets et qui contient les
« autres dans son intérieur? Toutes ces gran-
« deurs, dans le groupe des Pléïades, sont plus
« petites qu'il n'est vraisemblable. Laquelle
« d'entre elles donnera la mesure de l'invraisem-

1. Bertrand, *Calcul des Probabilités*, p. 170.

« blance? Si trois étoiles forment un triangle « équilatéral, faut-il faire entrer cette circons- « tance. assurément peu probable *a priori*, au « nombre de celles qui révèlent une cause? »

Ce que l'on traduit en chiffres, c'est la probabilité pour qu'un groupement soit dû au hasard; quoi qu'en dise Bertrand, le résultat numérique ainsi défini ne dépendra que dans une très faible mesure de la forme particulière donnée à la définition du groupement : plus petit cercle le renfermant, ou polygone sphérique convexe, etc. Pratiquement, on choisira celle de ces définitions qui conduira aux calculs les plus simples. Ces calculs conduiront à un résultat tel que le suivant : le hasard produirait une fois sur 23 000 un tel groupement. Une fois ce résultat énoncé, la théorie des probabilités a donné tout ce qu'elle pouvait donner et doit céder la place à l'induction et à l'hypothèse, que l'on cherchera ultérieurement à vérifier par d'autres méthodes, si c'est possible.

Disons un mot de la réflexion de Bertrand relativement au triangle équilatéral que formeraient trois étoiles; elle se rattache à la question du nombre rond. Si l'on considère un nombre pris au hasard entre 1.000.000 et 2.000.000 la probabilité pour qu'il soit égal à 1.342.517 est égale à un millionième; la probabilité pour qu'il soit égal à 1.500.000 est aussi égale à un millionième. On considérera cependant volontiers cette dernière éventualité comme moins probable que la première; cela tient à ce qu'on ne se représente jamais *individuellement* un nombre tel que 1,542.317; on le regarde comme le *type* de nombres d'apparences analogues et si, en le

transcrivant, on modifie un chiffre, on s'en aperçoit à peine et l'on ne distingue pas 1.324 519 de 1.324.517 : le lecteur a besoin de faire un effort pour s'assurer que les quatre nombres écrits dans les lignes précédentes sont tous différents.

Lorsque l'on a observé un nombre tel que le précédent comme évaluation d'un angle en dixièmes de secondes centésimales, on ne songe pas à se poser la question de savoir quelle était la probabilité pour que cet angle fût précisément égal à $13^{G}42^{\backprime}51^{\backprime\backprime}7$; car on ne se serait jamais posé cette question précise avant d'avoir mesuré l'angle. Il faut bien que cet angle ait une valeur et, quelle que soit sa valeur à un dixième de seconde près, on pourrait, après l'avoir mesurée, dire que la probabilité *a priori*, pour que cette valeur soit précisément telle qu'elle est, est un dix-millionième, et que c'est là un fait bien extraordinaire. C'est un sophisme de ce genre qui viciait un système d'expertise fondé sur la mensuration des écritures, système qui eut un moment de célébrité il y a quelques années, pour des raisons complètement étrangères aux mathématiques.

La question est de savoir si l'on doit faire ces mêmes réserves dans le cas où l'on constate qu'un des angles du triangle formé par trois étoiles a une valeur *remarquable* et est, par exemple, égal à l'angle du triangle équilatéral ($66^{G}66^{\backprime}66^{\backprime\backprime}6$), ou à un demi-angle droit : 50 grades, ou 5.000.000 dixièmes de seconde. Voici ce que l'on peut dire à ce sujet : on doit se défier beaucoup de la tendance que l'on a à regarder comme *remarquable* une circonstance que l'on n'avait pas précisée *avant l'expérience*, car le nombre des

circonstances qui peuvent apparaître comme remarquables, à divers points de vue, est très considérable[1].

Or la valeur de l'angle déterminé par les positions apparentes de trois étoiles est une fonction compliquée des positions réelles des étoiles et de la position de notre système solaire; le fait que cet angle serait précisément égal à 6 millions 666.666 dixièmes de seconde centésimale ne doit pas paraître plus remarquable que le fait qu'il serait égal précisément à 1.234.567 dixièmes de seconde, ce qui est aussi un nombre « remarquable », ou à 1.259.921, ce qui est un nombre ne présentant *a priori* rien de particulier.

Au contraire, la question du groupement dans l'espace est une de celles qu'il est naturel de se poser *a priori;* la théorie des probabilités ne permet pas de la résoudre avec certitude, mais précise les conséquences que l'on peut tirer des observations.

Cette précision n'est pas inutile, car elle remplace souvent d'une manière heureuse les induc-

1. Voici un exemple simple de ce fait; considérons un nombre de quatre chiffres tels que 2.545 et supposons qu'on le décompose par la pensée en deux tranches de deux chiffres 25 et 45 ; on remarque ici que ces deux nombres se terminent par le même chiffre 5, ce qui est une *circonstance particulière;* de même pour 2.524, les deux nombres 25 et 24 commencent par le même chiffre 2 ; pour 2.552 les nombres 25 et 52 sont symétriques ; pour 2.550 le second nombre 50 est double de 25, pour 2.536, 25 et 36 sont les carrés de deux entiers consécutifs, etc. On en arrive ainsi aisément à trouver à près de la moitié des nombres de quatre chiffres *quelque chose de particulier*, c'est là un petit jeu auquel on peut s'amuser en relevant les numéros de voitures publiques ou de wagons, dans les cas assez fréquents où ces numéros ont quatre chiffres ; on serait porté à croire, si l'on ne précisait pas, que les numéros *particuliers* sont plus fréquents que ne le voudrait la théorie des probabilités et à chercher la cause de cette fréquence.

tions trop rapides du « bon sens ». Sans doute, il n'est besoin d'aucun calcul pour savoir que la voie lactée est un amas exceptionnel d'étoiles, et le calcul n'ajouterait pas grand'chose à cette impression. D'autre part, il est des cas où l'on sera conduit à ce résultat que la probabilité d'un groupement est un tiers ou un quart; la présomption pour qu'il y ait une cause est ici trop faible pour que le résultat ait un intérêt : il ne prouve rien, pas plus que le fait qu'un joueur d'écarté a retourné une fois le roi ne prouve qu'il est un tricheur. Mais, entre ces deux cas extrêmes, celui où la probabilité du groupement par le hasard est tellement faible qu'on a l'intuition qu'elle est pratiquement nulle et celui où cette probabilité est assez élevée pour qu'on ne puisse rien conclure, il est bien d'autres cas, ceux où la probabilité est comprise entre un centième et un cent millième, par exemple : le calcul de la valeur qu'elle a entre ces limites n'est pas inutile et le renseignement qu'il fournit peut être précieux, quoi qu'en ait pu penser Bertrand.

43. — Je n'insisterai pas plus longuement sur les généralités relatives aux probabilités des causes; c'est en passant en revue les applications de la théorie des probabilités que nous nous rendrons compte de l'importance et de la variété des problèmes. Je signale cependant un cas très important, sur lequel je ne reviendrai pas, car il exige des développements mathématiques assez étendus, qui se trouvent dans un grand nombre d'ouvrages spéciaux : je veux parler de la théorie des erreurs d'observation. Plusieurs observateurs, mesurant une longueur, ont trouvé des

résultats légèrement différents entre eux; dans le langage que nous avons adopté, ces résultats ont pour cause commune la vraie longueur de la quantité mesurée; c'est cette cause qu'il s'agit de déterminer d'après ses effets, c'est-à-dire d'après les mesures. Indiquons simplement que l'on est conduit à considérer comme la valeur la plus probable de la quantité mesurée la moyenne arithmétique des mesures. Cette règle de la moyenne donne lieu d'ailleurs à des objections analogues à celles que nous avons étudiées à propos des probabilités continues, et ces objections se réfutent de la même manière, en se plaçant sur le terrain concret.

On a trouvé, en faisant deux mesures, les nombres 100 et 102, leur moyenne est 101. Mais on aurait pu se proposer de calculer le carré de la quantité à mesurer; tel serait le cas si la mesure avait porté sur une surface ayant la forme d'un carré, dont on désirerait connaître la superficie; le carré de 100 est 10.000 et le carré de 102 est 10.404; la moyenne est 10.202 et diffère du carré de 101, qui est 10.201. La différence est faible, il est vrai, et il en sera de même toutes les fois que les erreurs seront faibles, mais il suffit qu'elle existe pour que la règle de la moyenne ne puisse prétendre à la valeur absolue que lui attribuent ceux qui la déduisent de la théorie de Gauss.

Telle est l'objection. Elle ne me semble pas irréfutable; l'erreur de mesure est en effet indissolublement liée avec la technique même de la mesure, et cette technique atteint directement une quantité d'espèce bien déterminée en fait, ici, une longueur. S'il s'agissait du volume d'un

cube formé d'une substance homogène, de den sité connue, on pourrait l'atteindre de deux ma nières, en mesurant l'arête ou en pesant le cube dans le premier cas, la règle de la moyenn devrait s'appliquer aux longueurs mesurées, e dans le second cas, aux poids (ou aux volume qui leur sont proportionnels); il serait déraison nable de vouloir, avant de prendre la moyenne remplacer les longueurs par leurs cubes ou le volumes par leurs racines cubiques.



DEUXIÈME PARTIE

L'APPLICATION DES LOIS DU HASARD

CHAPITRE V

LES SCIENCES SOCIOLOGIQUES ET BIOLOGIQUES

44. — C'est dans les sciences les plus complexes que les lois du hasard ont été le plus aisément applicables et le plus rapidement fécondes. Il est donc à la fois conforme à l'état actuel des sciences et à l'ordre historique d'étudier les applications scientifiques des lois du hasard dans un ordre opposé à celui dans lequel Auguste Comte a rangé les sciences; tout d'abord, sociologie et biologie, ensuite sciences physiques et enfin sciences mathématiques.

45. — Les phénomènes sociaux les plus importants sont les phénomènes démographiques; il y a plus d'un siècle que dans les pays civilisés on a compris l'intérêt qui s'attache à l'étude statisti-

que précise de ces phénomènes ; ces statistiques sont plus ou moins détaillées suivant les pays ; la plupart des grandes villes sont à cet égard en avance sur le pays auquel elles appartiennent ; les naissances sont étudiées d'après l'âge des parents, les décès classés suivant l'âge, suivant le sexe, les professions, les causes de maladies, etc.

Certains auteurs distinguent les *rapports statistiques* que l'on peut calculer au moyen des résultats statistiques ainsi publiés, des *coefficients de probabilité* qui se présentent, par exemple, dans l'étude du jeu de pile ou face ou du jeu de dés. Cette distinction peut avoir l'utilité pratique de rappeler que les résultats de toute statistique comportent une certaine imprécision et que par suite les rapports déduits de ces résultats ne sont qu'approchés, en raison d'abord de l'imprécision des résultats, en raison ensuite du fait que les statistiques ne correspondent qu'à un nombre limité d'expériences, en raison enfin de cette circonstance que la matière statistique elle-même se modifie avec le temps. Mais, si l'on y réfléchit, on constatera que ces caractères se retrouvent à quelque degré dans tous les coefficients de probabilité définis d'une manière concrète[1]. Lorsque l'on admet que les six faces d'un dé ont le même coefficient de probabilité, égal par suite à un sixième pour chacune d'elles, on

1. Bien entendu, on peut, comme certains auteurs l'ont proposé, donner de la probabilité une définition purement abstraite, comme on fait de la circonférence en géométrie ou de la masse en mécanique rationnelle ; d'une telle définition, on pourra déduire des conséquences logiques absolument rigoureuses ; mais lorsqu'on voudra appliquer ces conséquences à un problème réel quelconque. il faudra remplacer la probabilité abstraite par la probabilité concrète d'un phénomène réel.

se réfère, soit à la précision supposée de la fabrication du dé, soit à des expériences en nombre limité faites sur ce dé ou d'autres dés supposés analogues ; on postule en outre que ce dé se modifie assez peu par l'usage pour que les coefficients de probabilité ne soient pas modifiés.

L'application des méthodes de la théorie des probabilités aux rapports statistiques est donc légitime ; il faut simplement ne pas oublier que les conséquences ne peuvent jamais avoir plus de précision que les prémisses, et qu'elles en ont souvent moins lorsqu'elles sont séparées de ces prémisses par des déductions compliquées (ou de longs calculs). Il convient donc, dans l'étude des phénomènes statistiques complexes, de n'user du calcul qu'avec modération, simplement pour préciser un peu les indications du bon sens. Sous cette forme modeste, l'étude des statistiques peut conduire à des conséquences sociales intéressantes, dont la plupart s'énoncent sous forme de *corrélations* (nous reviendrons sur cette notion de corrélation à la fin de ce chapitre). Par exemple, il y a corrélation directe entre le nombre des décès des enfants âgés de moins d'un an et le nombre des naissances l'année suivante : ces deux nombres varient dans le même sens.

46. — Ce sont surtout les statistiques de décès qui ont été étudiées d'une manière sérieuse, car c'est sur cette étude que repose toute l'organisation des assurances sur la vie, dont on n'ignore pas l'extension très considérable, on peut presque dire universelle, si l'on tient compte des lois qui organisent actuellement dans les pays civili-

sés l'assurance par l'État au profit de toute la classe des travailleurs. En se bornant aux entreprises privées d'assurances, le métier d'*actuaire* comporte l'étude de la mortalité pour la clientèle spéciale de ces entreprises, le calcul des primes qui résultent du taux de mortalité pour les risques simples ou les risques multiples, de la majoration à exiger pour obtenir un bénéfice déterminé, des risques courus en raison des fluctuations qui se produisent dans la mortalité par rapport aux nombres qui résulteraient de l'application stricte des coefficients de probabilité déduits des tables, l'établissement de la valeur du portefeuille d'une compagnie d'assurances à une époque donnée, la valeur à attribuer à un contrat en cas de résiliation, etc. L'actuariat est actuellement une des branches les plus importantes des mathématiques appliquées ; le développement récent des assurances de toute espèce ne peut qu'accroître cette importance ; il y a là un champ très étendu d'applications pour la théorie des probabilités. L'étude technique des méthodes particulières qui y sont employées ne présente guère d'intérêt que pour les spécialistes ; je préfère insister sur d'autres applications sociales de la théorie des probabilités, moins précises sans doute, moins riches surtout en dividendes pour les actionnaires, mais peut-être plus instructives en ce qui concerne le rôle de la théorie des probabilités dans les questions humaines.

47. — Il se présente dans bien des questions économiques un paradoxe que l'on peut rattacher à ce que l'on appelle dans les cours de logique le « sophisme du tas de blé ». Parmi les sophismes

que nous ont légués les Grecs, nul plus que ce « sophisme du tas de blé » n'a mérité de traverser les siècles ; ce n'est point en effet un pur jeu d'esprit, mais bien un exemple topique d'une difficulté fréquente, tant dans la vie pratique que dans la spéculation pure.

Un grain de blé ne constitue pas *un tas;* ni deux grains ; ni trois grains ;... d'autre part, si un million de grains sont réunis, on s'accordera pour dire qu'ils forment un tas. Quelle est la limite exacte ? Devra-t-on dire que 2.342 grains, par exemple, ne forment pas un tas, tandis que 2.343 grains en forment un ? C'est manifestement ridicule. On ne voit aucun moyen logique de sortir de l'impasse ; il n'est donc pas possible de savoir ce que c'est qu'un tas de blé.

On peut varier de bien des manières l'énoncé du sophisme ; une classification de ces énoncés ne sera pas inutile.

Une première catégorie est toute semblable au sophisme primitif : à quel moment de sa vie un enfant doit-il être qualifié d'homme? Quelle est la limite qui sépare une masure d'une maison, une maison d'un palais ? A quel instant une série de teintes savamment dégradées passe-t-elle du gris au blanc, ou du vert au bleu ? Une attitude possible vis-à-vis de ces questions consiste à les déclarer complètement dénuées d'intérêt, vu qu'il s'agit seulement d'une définition de mots, qu'on a le droit de regarder comme purement arbitraire. Nous dirons pour quelles raisons cette attitude ne nous paraît pas satisfaisante, mais il faut reconnaître que, si c'était là toute la portée du sophisme du tas, il ne mériterait peut-être pas une étude approfondie.

Une deuxième catégorie de questions paraîtra sans doute plus importante. Je veux parler des cas où la distinction à établir n'est pas une simple distinction verbale, mais entraîne des conséquences pratiques. Les administrations résolvent en général la question par une décision arbitraire, mais simple. Telle était le minimum de taille nécessaire pour servir dans l'armée, supprimé depuis peu ; tel est encore l'âge de trois ans au-dessous duquel les enfants voyagent gratuitement dans nos chemins de fer. Mais lorsqu'on a à prendre une décision personnelle pour laquelle on n'a pas prévu de règles administratives, on se trouve quelquefois en face de difficultés de ce genre ; on veut louer une chambre à coucher et l'on attache une importance particulière à ce que le cube d'air soit suffisant ; on est bien décidé à accepter une chambre de 100 mètres cubes et à refuser une chambre de 10 mètres cubes ; quelle est la limite précise que l'on fixera ? Et quelle que soit cette limite, qui ne consentira à l'abaisser d'une fraction de mètre cube si les autres conditions sont avantageuses et agréables ?

Enfin, dans une troisième catégorie, nous rangerons des questions qui sont objectivement analogues aux précédentes, mais qui en diffèrent subjectivement en ce que, au lieu d'avoir à prendre une décision personnelle au sujet du sophisme, nous devons prévoir les décisions que prendront un grand nombre de personnes, décisions dont l'ensemble a pour nous une grande importance. C'est le cas de tout commerçant fixant des prix de vente, de toute entreprise de transports modifiant ses tarifs, etc. La même difficulté se pré-

sente dans bien des questions de fiscalité, compliquée par la discontinuité monétaire ; on augmente de un franc par hectolitre les droits d'octroi sur la bière ou le vin ; le prix du litre se trouve donc augmenté de *un centime* et le prix du quart de litre de *un quart de centime*. Le prix de vente au détail, prix qui ne peut varier que par multiples de cinq centimes, sera-t-il modifié ? C'est en étudiant cette troisième catégorie que nous éluciderons le paradoxe économique dont nous voulons parler, et dont l'énoncé ressort déjà clairement des exemples que nous venons de donner.

Il serait facile de multiplier ces exemples ; le lecteur en imaginera sans peine un très grand nombre ; mais ceux qui précèdent suffisent sans doute pour mettre en évidence la variété et l'importance des questions qui se rattachent au sophisme du tas de blé ; je voudrais essayer de montrer que l'on peut donner à toutes ces questions une réponse parfaitement claire, si l'on y fait systématiquement usage des principes de la théorie des probabilités. Ce sont en effet des questions mal posées, si l'on exige une réponse par oui ou par non ; la vraie réponse est un coefficient de probabilité.

48. — Reprenons d'abord les questions les plus simples, où la difficulté paraît être purement verbale. Elle le serait, en effet, s'il s'agissait de créer une langue artificielle. On pourrait évidemment décider que, dans un certain volapük, les mots « tas de blé » désignent la réunion de 1.000 grains, mais ne s'appliquent pas à la réunion de 999 grains. Une telle décision serait arbitraire,

mais légitime pour qui admet la légitimité de la création d'un volapük[1].

Mais il ne s'agit point de la langue volapük ; c'est dans la langue française qu'il s'agit de préciser le sens des mots « tas de blé », et le français n'est pas une langue arbitraire, dont nous puissions disposer à notre gré. C'est une langue qui nous est donnée comme un fait. Il est d'ailleurs possible de discuter sur la nature précise de ce fait; les uns regardent le français comme défini par les « bons » écrivains, les autres par la conversation des gens « cultivés » ; d'autres enfin prendront un critérium plus démocratique. Mais, quelle que soit la définition adoptée, à moins qu'on ne se réfère à un dictionnaire parfait (ce qui ne serait possible que s'il existait un tel dictionnaire), on est obligé, pour définir le sens précis d'un mot, de faire appel à de nombreux témoignages. C'est d'ailleurs par de nombreux exemples que les meilleurs dictionnaires tâchent de fixer le sens des mots : c'est leur seule chance d'échapper au cercle vicieux fondamental qui consiste à définir des mots par des mots.

S'il s'agit donc de définir un terme usuel, tel que celui de maison ou de masure, on sera sans doute d'accord pour accepter le témoignage d'un certain nombre de Français, sinon de tous. A ce point de vue, la solution du sophisme est facile. Qu'on imagine mille Parisiens défilant devant un immeuble à sept étages de l'avenue de l'Opéra ;

1. J'ai à peine besoin de dire que je ne prétends pas résoudre en passant la question si discutée de la création possible d'une langue artificielle viable. Il me paraît simplement évident que, du jour où une telle langue serait viable, elle aurait précisément, au point de vue qui nous intéresse ici, les caractères d'une langue naturelle.

ils seront tous d'accord pour le qualifier de maison, tandis qu'ils refuseront tous ce titre à une construction de pierre pouvant abriter deux lapins et trois poules. Qu'il s'agisse maintenant d'un édifice intermédiaire, les avis pourront être partagés ; si 748 opinants sur 1.000 dénomment l'édifice maison, il sera raisonnable de convenir de dire que la probabilité pour qu'il soit une maison est 0,748, la probabilité contraire étant 0,252. Cette manière de s'exprimer paraîtra singulière à bien des gens : à la réflexion, on se convaincra que c'est la seule réponse raisonnable à faire à l'argument de Zénon.

Il est d'ailleurs clair que si, au lieu de consulter 1.000 Parisiens, on avait consulté 1.000 habitants d'une petite ville, puis 1.000 paysans, la proportion des réponses eût été différente ; il est nécessaire de faire une convention précise pour que la définition que nous avons donnée ne prête pas à ambiguïté ; mais ces difficultés d'application n'ont rien à voir avec le principe de la théorie ; elles sont relatives à l'indétermination forcée des définitions verbales ; aussi avons-nous hâte d'arriver à des questions d'un intérêt plus réel, nous contentant d'avoir attiré l'attention sur ce fait que, dans les questions de langage comme dans beaucoup d'autres, *il n'y a que des vérités statistiques.*

49. — Un éditeur entreprend une collection d'ouvrages, qui, pour des raisons commerciales, doivent être à peu près tous de la même dimension : 251 à 300 pages d'un type bien déterminé. Si donc on lui propose un manuscrit devant donner 50 pages de ce type, ou devant en donner 500,

il n'hésitera pas à le refuser. Mais la difficulté du tas de blé se présentera pour les cas intermédiaires. Si, pour préciser, nous admettons que le type adopté comporte en moyenne 1.000 caractères à la page, les dimensions extrêmes correspondent à un manuscrit de 250.001 à 300.000 caractères ; faudra-t-il refuser un manuscrit qui en renfermerait 250.000, ou 249.999, ou 249.998... : c'est le problème du tas de blé.

On peut imaginer bien des manières de résoudre pratiquement la question : il ne sera pas inutile d'en énumérer quelques-unes, car si la difficulté essentielle reste la même, elle se présentera sous des formes diverses et instructives.

Une première solution est « administrative ». Avant de lire le manuscrit, l'éditeur le fait « évaluer » par un de ses employés ; celui-ci doit lui indiquer le nombre exact de pages qu'il suppose devoir être fournies par le manuscrit ; si l'évaluation est inférieure à 251 ou supérieure à 300, le manuscrit ne sera pas accepté, du moins sous sa forme actuelle. Nous supposons, bien entendu, que l'employé chargé de l'évaluation est consciencieux et incorruptible ; mais il serait contraire au bon sens de le supposer infaillible. Quelle est alors la situation d'un auteur dont le manuscrit compte 249.999 caractères, ou 250.000, ou 250.001 ? Si l'employé chargé de l'évaluation est très habile, *il y a une chance sur deux* pour que le manuscrit soit accepté, telle est la réponse précise que l'on peut faire. Car, si habile qu'il soit, il ne pourra pas éviter de commettre une erreur de quelques unités (et même de quelques centaines d'unités), en plus ou en moins ; le supposer très habile ne peut donc avoir qu'un sens : *il y a*

autant de chances pour qu'il se trompe en plus que pour qu'il se trompe en moins ; c'est précisément ce que nous exprimons en disant qu'il y a une chance sur deux pour que son évaluation soit favorable à l'auteur.

Pour continuer l'étude de la question, il faudrait faire intervenir l'expression mathématique de la loi des écarts et supposer que l'on connaît l'écart moyen (c'est-à-dire l'erreur moyenne commise dans son évaluation par l'employé ; cette erreur dépend évidemment à la fois de l'habileté de l'employé et de la régularité plus ou moins grande du manuscrit). Je n'entrerai pas dans le détail des formules ; je me contenterai d'énoncer le résultat évident sans formules pour quiconque a suivi nos raisonnements : lorsque le nombre des caractères du manuscrit décroît unité par unité, on peut calculer la probabilité pour qu'il soit évalué comme acceptable, et par suite accepté : cette probabilité décroît lentement, d'une manière continue, et finit par devenir sensiblement égale à zéro lorsque le nombre des caractères est descendu au-dessous d'une certaine limite, qui dépend de l'habileté de l'employé. En d'autres termes, sur 1.000 manuscrits de 250.000 caractères, 500 seront acceptés ; sur 1.000 manuscrits de 249.000 caractères, 450 peut-être seulement sont acceptés..., sur 1.000 manuscrits de 230.000 caractères, un seul peut-être sera accepté. La discontinuité qui répugnait à l'esprit disparaît dans la pratique : mais il faut admettre pour cela que, dans les cas douteux, la seule réponse raisonnable est un coefficient de probabilité.

Après la solution « administrative », disons quelques mots de la solution « esthétique ». Il

s'agit d'un recueil de vers, et la disposition typographique adoptée entraîne la connaissance exacte du nombre de pages ; il est d'ailleurs possible de « tricher » en intercalant un certain nombre de pages blanches, avec des titres et sous-titres variés. Pour obtenir un volume présentable, l'éditeur juge qu'il faut au moins 120 pages « utiles »; le manuscrit ne lui en donne que 119 ; que fera-t-il ? La réponse est encore ici un coefficient de probabilité, mais il s'y mêle un élément d'appréciation littéraire. Suivant l'opinion qu'il a de la valeur du texte ou de la notoriété de l'auteur, l'éditeur sera plus ou moins exigeant sur la quantité de matière ; mais la probabilité pour qu'il accepte un nombre donné de pages décroîtra lentement en même temps que ce nombre diminuera ; ici encore il n'y aura pas de saut brusque, mais décroissance insensible jusqu'à la probabilité zéro : aucun éditeur n'aurait consenti à publier quatre lignes de Victor Hugo en un volume de 3 fr. 50.

On pourrait aussi imaginer une solution « mécanique » ; cette solution serait la plus naturelle si les auteurs avaient l'habitude de donner à leurs manuscrits la forme de ces feuilles de papier perforé, qu'il suffit de confier à une machine à composer pour que le livre soit cliché mécaniquement, la besogne de l'ouvrier étant seulement de surveillance. Mais, pour ne pas anticiper et rester dans la réalité présente, prenons un autre exemple. Un tamis mécanique soigneusement exécuté ne laisse passer que les grains de sable (ou de farine) qui sont au-dessous d'une certaine dimension. Si nous désignons cette dimension par 100, il est clair qu'un grain de dimension 50 passera toujours et qu'un grain de dimension 150

ne passera jamais ; mais si la dimension du grain varie lentement au voisinage de 100, il n'y aura certainement pas de discontinuité brusque ; si l'on suppose que les grains de dimension 100 passent tous, il passera en outre une fraction notable des grains de dimension 101, etc. La machine procède au fond, exactement de la même manière que l'homme ; seul, l'écart moyen peut varier suivant la plus ou moins grande perfection du tamis, mais dans tous les cas, ce sont les lois du calcul des probabilités qui régissent les variations : il n'y a jamais démarcation absolue, mais toujours dégradation insensible.

Il est inutile de multiplier les exemples ; tous nous conduiront à la même conclusion : dans toute question dont la solution exige que nous fixions « le nombre de grains de blé qui forment un tas », nous sommes dans l'impossibilité de donner une réponse autrement qu'en examinant un grand nombre de cas et en interrogeant les moyennes ; le calcul des probabilités seul nous permet de sortir de l'impasse logique où nous accule le raisonnement par continuité.

On peut d'ailleurs distinguer deux formes différentes de la question ; dans les exemples que nous avons examinés, nous avons supposé que l'on s'est fixé *a priori* une règle ; mais les faits ne s'assujettissent pas exactement à la règle, qui se trouve n'avoir qu'une valeur relative [1] : elle

1. Il en est ainsi, même dans les circonstances où l'on a cherché à écarter le plus possible les cas douteux par une réglementation administrative. Je ne parle pas de la gratuité de parcours sur les chemins de fer pour les enfants de moins de trois ans : car peu de parents ont les scrupules du père de famille dont un humoriste rapporte qu'il tirait la sonnette d'alarme à la seconde précise où son enfant atteignait l'âge de trois ans, afin de payer sa place pour le

détermine la moyenne. Dans d'autres cas, il pourra arriver que l'on ne se fixe pas de règle *a priori*, mais que l'on déduise la règle moyenne des faits eux-mêmes. Par exemple, lorsque l'on veut acheter un meuble, on n'a pas toujours une idée absolument arrêtée des dimensions que l'on désire ; cependant, il est certains modèles que l'on écartera sûrement comme trop petits, et certains que l'on écartera sûrement comme trop grands. En fait, si l'on voit beaucoup de modèles et que l'on note ceux dont les dimensions paraissent acceptables, on constatera que ces dimensions se répartissent autour d'une certaine moyenne, suivant des lois déjà rappelées. Dans un lot de modèles de même dimension, que l'on examinera à des instants différents, on en regardera comme acceptable une proportion d'autant plus faible que la dimension commune s'écartera davantage de la moyenne.

50. — Arrivons enfin au cas le plus fréquent et le plus important, et qui, pour ces motifs, est souvent discuté, mais presque toujours sans aucun esprit scientifique. Il s'agit des répercussions économiques de variations de prix très faibles par rapport à l'unité monétaire. L'un des exemples les plus caractéristiques est l'influence du prix de vente en gros sur le prix de vente au dé-

nombre de kilomètres restant à parcourir. Prenons le cas du minimum de taille qui fut exigé pour être déclaré apte au service militaire ; il est clair que, pour celui dont la taille différait de quelques millimètres en plus ou en moins de ce minimum, il y avait, quelque consciencieux que l'on supposât le conseil de revision, seulement une probabilité plus ou moins grande d'être accepté ou réformé. En d'autres termes, un même conscrit se présentant devant 100 conseils de révision différents, aurait été accepté par exemple 55 fois et réformé 45 fois.

tail. C'est presque un axiome pour certains publicistes que si le prix du gros diminue très peu, cette diminution correspondant à une fraction de centime dans le prix du détail ne pourra pas être réalisée pratiquement, c'est-à-dire que *le prix du détail ne variera pas*. Lorsque cette thèse est développée dans un article de journal, elle est souvent accompagnée de considérations sur les avantages que présenterait l'usage courant des pièces de un centime. Mais tenons-nous-en à l'affirmation elle-même. Son absurdité ressort avec évidence de l'argument du tas de blé : si le prix du gros baisse chaque semaine d'une même quantité, correspondant à un demi-centime sur le prix du détail, au bout de l'année la baisse du prix de détail devra atteindre 26 centimes ; il est bien certain que les plus pessimistes accorderont que cette baisse aura été, au moins partiellement, réalisée. Il y aura donc eu au moins une semaine dans l'année où la diminution théorique de un demi-centime aura entraîné une diminution pratique de cinq centimes.

Avant d'aller plus loin, il est nécessaire de répondre à une objection qui se présentera certainement à certains esprits. « Sans doute, dira-t-on, si les diminutions de un demi-centime se succèdent, il arrivera bien un moment où leur accumulation déterminera une baisse de cinq centimes ; mais c'est seulement lorsque le total des diminutions successives aura atteint effectivement cinq centimes. Il reste certain qu'une diminution d'un demi-centime est sans effet ; dix diminutions de un demi-centime produisent seules un résultat appréciable. »

L'argument est spécieux, mais il suffit de l'exa-

miner d'un peu près pour se convaincre qu'il est entièrement vide. Ce qui importe, c'est l'effet produit, en une semaine déterminée, par la diminution de un demi-centime qui se produit cette semaine-là : cet effet est une baisse de cinq centimes. Que cette baisse de cinq centimes ne soit pas due *uniquement* à la diminution de un demi-centime, c'est ce que personne ne songe à contester ; elle a certainement beaucoup d'autres causes, parmi lesquelles les diminutions antérieures doivent être spécialement mentionnées ; mais le fait simple qui n'est pas contestable est le suivant : la diminution de un demi-centime a été, *dans des conditions déterminées*, la cause immédiate de la baisse de cinq centimes ; si la diminution ne s'était pas produite, la baisse n'aurait pas eu lieu. Cette constatation nous suffit. Il n'est pas nécessaire, en effet, de pousser bien loin l'analyse pour se rendre compte que ces *conditions déterminées*, dans lesquelles une diminution de un demi-centime produira une baisse de cinq centimes *se produiront précisément une fois sur dix* ; en d'autres termes, la probabilité pour que la baisse se produise est un dixième, lorsque la diminution est précisément le dixième de l'unité monétaire qui fixe le minimum de la baisse possible.

Cet énoncé n'est d'ailleurs pas purement théorique ; en pratique, dans une grande ville ou dans un pays, une même marchandise n'est pas vendue partout le même prix, pour des raisons très diverses ; il arrivera donc que la baisse se produira dans la dixième partie environ des magasins de détail[1].

1. Bien que je ne puisse entrer ici dans tous les détails de cet exemple économique, il est nécessaire que je prévienne un malen-

Il se passe un phénomène analogue lorsque le prix d'une marchandise diminue brusquement ; une diminution, même légère, augmente le nombre des acheteurs ; et l'on pourrait dans certains cas (lorsqu'il s'agit d'un train de plaisir, du prix d'entrée à un spectacle, etc.) établir la loi d'après laquelle le nombre des clients augmente lorsque le prix diminue. Si le nombre total des clients est très considérable, on constate que l'augmentation de ce nombre est sensible, même pour une diminution insignifiante, de cinq centimes sur un prix de plusieurs francs. Et cependant, chaque personne prise individuellement est convaincue que ce n'est pas un sou de plus à payer qui l'aurait empêchée de faire cette excursion et d'assister à ce spectacle. L'illusion s'explique naturellement ; pour chacun, la probabilité de se décider diminue lentement, d'une manière continue, à mesure que le prix augmente ; mais, une fois la décision prise, on s'imagine volontiers que la probabilité que l'on avait de se décider était égale à l'unité, c'est-à-dire était la

tendu. Je me place dans le cas où la *qualité* de l'objet vendu ne peut pas être modifiée d'une manière continue (vin additionné d'eau, etc.). Il faut alors admettre que le détaillant fixe le prix de vente en tenant compte à la fois du prix de revient, du bénéfice qu'il désire et des exigences de la clientèle. Il arrive donc souvent qu'il hésite entre deux prix, par suite il suffira quelquefois d'une très légère différence dans les conditions de son hésitation pour faire pencher la balance d'un côté ou de l'autre. Il ne faut pas répondre à cela que le détaillant vend toujours « le plus cher possible » et que, comme on dit, l'acheteur est toujours « mis dedans » ; je ne fais point ici un plaidoyer en faveur du commerce au détail qui prélève certainement un impôt hors de proportion avec les services qu'il rend ; mais il faut reconnaître que les mots « le plus cher possible » n'ont aucun sens précis ; les bénéfices possibles ne sont pas illimités ; et la discontinuité monétaire empêche qu'ils soient constants ; ils ne peuvent qu'osciller autour d'une moyenne ; il est donc nécessaire qu'ils diminuent parfois.

certitude. Dès lors, si l'on a l'impression qu'une circonstance déterminée aurait diminué légèrement cette probabilité que l'on suppose égale à l'unité, on se figure qu'elle serait restée assez voisine de la certitude, c'est-à-dire pratiquement identique à la certitude. Cette illusion psychologique provient donc de ce que la décision prise réagit sur l'individu et acquiert rapidement une force qu'elle n'avait pas réellement. Ce genre d'illusion est fréquent ; notre esprit est souvent en équilibre instable : il suffit d'une très petite action extérieure pour rompre l'équilibre, et cette rupture d'équilibre entraîne des conséquences importantes, dont l'importance même masque à nos yeux le fait que l'équilibre était instable et que la rupture d'équilibre aurait pu se produire dans une direction opposée [1]. Lorsque l'on regarde les choses du dehors et non plus du dedans, on peut observer effectivement le phénomène de rupture d'équilibre, sans que l'illusion psychologique intervienne : on peut ainsi se rendre compte par la statistique de l'effet produit par une très faible variation de la probabilité, lorsqu'elle s'applique à un grand nombre de cas. Je laisse de côté la question du privilège des chiffres ronds ; la baisse de 8 francs à 7 fr. 95 est psychologiquement plus importante que la baisse de 8 fr. 20 à 8 fr. 15.

Des considérations analogues s'appliqueraient aux salaires et à bien d'autres questions, mais nous devons nous borner et conclure.

1. Ceci peut être rapproché de la définition qu'a donnée Poincaré du hasard : il y a hasard lorsque de petites causes produisent de grands effets. La cause insignifiante n'est pas perçue, mais l'effet nous frappe et nous paraît fortuit. Voir la *Revue du Mois* du 10 mars 1907, t. III, p. 257.

51. — L'idée que je voudrais dégager de ce qui précède, c'est que la réponse mathématique à donner à bien des questions pratiques est un *coefficient de probabilité.* Une telle réponse ne paraîtra pas satisfaisante à bien des esprits, qui attendent des mathématiques la certitude. C'est là une tendance très fâcheuse ; il est extrêmement regrettable que l'éducation du public soit, à ce point de vue, si peu avancée ; cela tient sans doute à ce que la théorie des probabilités est à peu près universellement ignorée, bien qu'elle pénètre chaque jour davantage dans la vie de chacun (assurances diverses, mutualités, retraites, etc.). Un coefficient de probabilité constitue une réponse tout à fait claire, correspondant à une réalité absolument tangible. Certains esprits maintiendront qu'ils « préfèrent » la certitude ; ils « préféreraient » peut-être aussi que 2 et 2 fissent 5.

Si la notion de *vérité statistique* devenait familière à tous ceux qui parlent ou écrivent au sujet de questions où la vérité statistique est la seule vérité, bien des sophismes et bien des paradoxes seraient évités. Beaucoup d'esprits, par ailleurs excellents, s'imaginent qu'il n'est de vérités que des vérités particulières, auxquelles ils donnent le nom de faits : *en débarquant à Douvres, j'ai aperçu trois Anglais dont la taille dépassait deux mètres*, ceci est un fait ; mais c'est un fait sans intérêt, tandis que la proposition suivante : *la taille moyenne des Anglais est inférieure à deux mètres*, n'exprime pas *un fait*, mais la moyenne d'un ensemble de faits, et c'est vraiment une vérité scientifique.

52. — C'est Quételet qui, au milieu du XIXe siècle, a eu le grand mérite d'attirer l'attention sur l'intérêt scientifique que pouvait présenter la comparaison, au point de vue de la théorie des probabilités, des mesures effectuées sur un groupe d'individus. Les recherches de Quételet sur la race humaine sont en quelque sorte un stade intermédiaire entre les applications sociologiques du calcul des probabilités et ses applications biologiques.

Supposons que l'on mesure la taille d'un très grand nombre de Français adultes ; si les mesures sont faites à un centimètre près, on pourra traduire leurs résultats globaux en disant combien il s'en est trouvé dont la taille est 1m,50, combien dont la taille est 1m,51, 1m,52, etc. Si l'on porte en abscisse les tailles et que l'on élève des ordonnées égales au nombre d'individus possédant cette taille, les extrémités de ces ordonnées décrivent une courbe vulgairement connue sous le nom de courbe en cloche, présentant un maximum unique et décroissant d'abord assez lentement, puis très rapidement à partir de ce maximum [1]. On donne aussi à ces courbes le nom de courbes binomiales, dont voici la signification. Considérons les coefficients du binôme de Newton pour une valeur donnée de l'exposant n

$$(a+b)^n = a^n + C_n^1 a^{n-1} b + C_n^2 a^{n-2} b^2 + \dots$$

et considérons une courbe (ou plutôt un polygone) tel qu'à l'abscisse p corresponde l'ordonnée C_n^p ;

1. L'équation de cette courbe est de la forme

$$y = ce^{-k(x-x_0)^2}$$

x_0 étant la valeur de x qui correspond au maximum.

ce polygone sera dit binomial ; si l'on choisit convenablement les unités de longueur pour les abscisses et les ordonnées, le polygone binomial a pour limite, lorsque n augmente indéfiniment, la courbe binomiale.

D'une manière plus simple, on peut dire que les tailles mesurées satisfont aux mêmes lois que les erreurs de mesure ; tout se passe comme si un même homme, dont la taille serait égale à la moyenne, avait été mesuré un grand nombre de fois par des observateurs assez maladroits ou ne disposant que d'instruments de mesure très imparfaits [1].

Quételet a déduit de ces remarques fort importantes des considérations dont quelques-unes sont critiquables. Il est néanmoins fort injuste de ne voir que les exagérations ou les erreurs de sa théorie de l'*homme moyen ;* il y a, dans les idées de Quételet, le germe des recherches biométriques qui ont pris une grande extension entre les mains de Francis Galton, de Karl Pearson et de leurs élèves.

53. — Signalons cependant une critique de Joseph Bertrand ; l'homme dont la taille est égale à la taille moyenne, le poids au poids moyen, etc., peut-il exister ? N'est-il pas un monstre ? Voici deux sphères dont l'une a pour rayon 1 et la seconde pour rayon 3 ; si elles sont

1. On peut dire aussi que tout se passe comme si à un jeu convenablement choisi une série de parties devait amener un nombre de parties gagnées égal au nombre de centimètres qui exprime la taille moyenne ; si l'on jouait autant de séries de parties analogues qu'il y a d'hommes mesurés, le tableau sur lequel on inscrirait le nombre de parties gagnées dans chaque série serait absolument comparable au tableau sur lequel seraient inscrites les tailles.

formées de même matière et si la première pèse un gramme, la seconde en pèsera 27. La « sphère moyenne » devrait avoir pour rayon 2, moyenne entre 1 et 3 et pour poids 14, moyenne entre 1 et 27. Mais, si elle est de même substance que les deux premières, au rayon 2 doit correspondre, non le poids 14, mais le poids 8. La géométrie ne permet aucune tricherie ; la biologie sera-t-elle moins exigeante ?

L'objection vaut d'être retenue ; on observera cependant que sa portée pratique est, en fait, moins grande que sa portée théorique.

On donne le nom de séries *normales* aux séries de mesures qui se répartissent suivant la loi dite loi de Gauss (courbe en cloche ou courbes binomiales) ; cette notion joue un rôle très important dans les recherches biométriques dont nous dirons un mot tout à l'heure. Mais auparavant, je voudrais insister un peu sur l'importance biologique de la notion de vérité statistique.

54. — On a célébré récemment le cinquantenaire de l'*Origine des Espèces* (1857), et le centenaire de la *Philosophie zoologique* (1809) ; en même temps, on a beaucoup parlé de la crise du transformisme. Mais cette crise est une crise de croissance ; il eût été étrange que les idées de Lamark et de Darwin échappent à la loi naturelle de l'évolution. Même si j'avais la compétence nécessaire, je n'aurais pas à discuter ici le détail des modifications proposées aux théories de l'évolution des êtres vivants [1]. Je voudrais

1. Voir l'excellent livre d'Étienne Rabaud : *le Transformisme et l'Expérience* (Nouvelle collection scientifique, Paris, F. Alcan) et l'ouvrage suggestif de Le Dantec : *la Crise du Transformisme* (Ibid.).

simplement, d'un point de vue extérieur, constater l'importance croissante que prennent les recherches expérimentales sur les espèces animales et végétales et chercher quelles conclusions peuvent se dégager de l'ensemble de ces recherches et des méthodes employées.

Observons tout d'abord qu'il y a, du moins en apparence, une contradiction logique entre la notion d'espèce et la notion d'évolution. Une définition logique de l'espèce doit en effet, semble-t-il, postuler la stabilité ; un certain groupe[1] d'animaux ne mérite d'être classé à part que si tous les animaux de ce groupe possèdent certains caractères communs, et si ces caractères appartiennent aussi à leurs descendants. Il est assez difficile d'étudier les caractères considérés comme fondamentaux par les biologistes sans entrer dans des détails techniques et sans faire usage de figures qui n'ont pas leur place ici ; il nous sera plus commode de raisonner sur un caractère simple connu de tous, la taille par exemple. On sait que certaines races humaines sont de petite taille, comparativement à la race blanche. Dans quelle mesure la taille peut-elle être considérée comme un caractère distinctif de la race ? Supposons que nous groupions dans une région isolée, telle qu'une île éloignée, des adultes de race blanche ayant très exactement la même taille[2], 1m,65 par exemple. Il est bien certain que dans trente ou quarante ans, la nouvelle population

1. J'emploierai cette expression vague de *groupe* afin de ne pas avoir à discuter la notion d'espèce, en opposition avec la notion de variété.

2. Je fais abstraction ici, pour simplifier le discours, du fait que les hommes sont généralement de taille plus élevée que les femmes.

adulte de l'île aura perdu ce caractère singulier ; la taille de certains individus sera supérieure, la taille de certains autres sera inférieure à 1m, 65 ; l'écart en plus ou en moins sera souvent de dix centimètres, peut-être même de quinze ou vingt centimètres. Nous voyons en effet tous les jours des parents tous deux de taille moyenne avoir des enfants de taille notablement inférieure ou notablement supérieure à la moyenne.

L'examen de ces premiers faits conduirait donc à cette conclusion : la taille de l'homme peut, d'une génération à la suivante, éprouver des variations en plus ou en moins, qui atteignent assez fréquemment une vingtaine de centimètres. Mais cette conclusion est évidemment inadmissible, sous cette forme du moins. Car, si une génération peut suffire à produire une variation de taille de vingt centimètres, deux générations produiraient une variation de quarante centimètres,... cinq générations une variation d'un mètre ; une variation d'un mètre par siècle conduirait à dix mètres en dix siècles : nous nageons dans l'absurde. Le lecteur a déjà aperçu le défaut de nos raisonnements : nous avons confondu la variation individuelle avec la variation collective. Il est bien exact que la taille d'un enfant devenu adulte peut dépasser notablement la taille de chacun de ses parents, mais ce fait ne se produira que si la taille des parents est relativement peu élevée, par rapport à la taille de la race. On trouvera assez aisément des exemples de parents mesurant 1m, 60 dont le fils mesure 1m, 80, on en trouvera plus difficilement de parents mesurant 1m, 80, dont le fils mesure 2 mètres ; les tailles plus élevées encore seront absolument exception-

nelles. De même pour les très petites tailles : le gigantisme et le nanisme sont manifestement des phénomènes pathologiques, au même titre que certaines difformités exhibées aussi dans les baraques foraines.

On peut résumer ces faits d'observation courante en disant que, si la taille des enfants dépend dans une certaine mesure de la taille des parents, elle dépend bien davantage de ce que nous avons appelé la taille de la race. Si les parents sont tous deux très grands, c'est-à-dire dépassent de beaucoup la taille de la race, il arrivera le plus souvent que les enfants seront plus petits qu'eux, c'est-à-dire se rapprocheront de la taille de la race. De même, si les parents sont très petits, leurs enfants seront généralement plus grands qu'eux. La taille moyenne à laquelle nous pensons lorsque nous disons qu'un homme est très grand ou est très petit n'est donc pas seulement une conception abstraite de notre esprit ; c'est une réalité, c'est même un caractère appartenant à chaque individu, au même titre que sa propre taille. Nous ignorons en quoi consiste d'une manière précise ce caractère ; nous ne serions pas capables, en étudiant un seul représentant d'une race inconnue, de déterminer la taille moyenne de cette race, ni même de dire si le représentant étudié est, par rapport à sa race, de grande ou de petite taille. Mais ce caractère n'en existe pas moins et se manifeste dans le cas où l'individu considéré a une descendance ; cette descendance tend à se rapprocher de la taille de la race.

55. — Tout cela est bien banal, et je m'excuse de m'y être attardé ; mais peut-être n'était-il pas

inutile d'insister sur ce fait que la notion de la taille moyenne de notre race est pour nous une notion commune ; car nous sommes ainsi encouragés à surmonter les difficultés qui se présentent lorsqu'on veut donner à cette notion une précision scientifique rigoureuse. Ces difficultés sont, au premier abord, inextricables. Veut-on définir la taille moyenne des Français adultes ? Il faudra d'abord préciser ce que l'on entend par un Français adulte : acceptons, puisqu'il le faut bien, une définition arbitraire : il s'agira des hommes nés en France continentale et y habitant, âgés de vingt-cinq à trente ans, par exemple ; on choisira un jour déterminé, et on mesurera toutes les tailles, ce même jour, à midi. On exclut donc ceux qui habitent l'étranger ou qui y séjournent ; que faire alors de ceux qui passent la frontière à midi ? ou qui naviguent près des côtes ? ou de ceux qui accomplissent ce jour-là leur vingt-cinquième ou trentième année, ou de ceux qui sont malades ou mourants ? Les difficultés seraient aussi grandes, si l'on prenait en considération la nationalité : que penser de ceux pour lesquels elle est contestée en ce moment même, devant les tribunaux ? Si l'on prend la taille des conscrits, tiendra-t-on compte des insoumis, des malades qui ne peuvent venir au Conseil de révision ?

Ces objections sont logiquement insurmontables, et cependant le sens commun ne peut s'empêcher de les juger sans portée, presque puériles. C'est le sens commun qui est dans le vrai contre la logique abstraite. Il est dans le vrai tout d'abord pour une raison de fait. La taille d'un individu n'est pas un nombre arithmétique, tel que la racine carrée de 2, qu'un mathématicien

peut calculer, s'il lui plaît et s'il en a le loisir, avec 10, avec 100, avec un million de décimales exactes ; c'est un nombre physique, difficile à mesurer avec une bien grande précision ; il varie d'ailleurs, pour un même individu, suivant les circonstances physiologiques (fatigue, marche, repos prolongé,...) et, tout ce que l'on peut espérer, c'est de le déterminer à un millimètre près ; prétendre à une plus grande précision n'aurait aucun sens, de par la nature même de l'objet de la mesure. Il en est forcément de même pour la taille moyenne, que l'on obtient en prenant la moyenne des mesures individuelles ; elle participe de leur incertitude, et n'est connue, elle aussi, qu'approximativement. Dans ces conditions, il est aisé de se rendre compte, sans que nous ayons à entrer dans le détail des calculs, que les possibilités d'erreurs dues aux imprécisions nécessaires de la définition sont plus faibles que les erreurs qui proviennent de la nature même de la question. En d'autres termes, le résultat final est le même, quelle que soit la solution adoptée pour les cas litigieux dont nous avons cité quelques exemples. On peut donc, en fait, négliger l'objection logique tirée de ces cas litigieux.

Quelle que soit la force de cet argument de fait, il ne satisfera peut-être pas tous les esprits. On peut se demander, en effet, si un perfectionnement des méthodes de mesure, et surtout l'emploi d'une technique physiologique précisant la définition de la longueur à mesurer, ne permettraient pas de calculer la taille moyenne avec une précision beaucoup plus grande, de sorte que les objections écartées reprendraient leur valeur : les cas litigieux auraient une influence

minime, il est vrai, mais accessible à la mesure. Devrons-nous pour cela renoncer à parler de la taille moyenne ? Evidemment non, seulement nous devrons dire que cette taille ne nous est connue qu'à un millième de millimètre près, par exemple. Cela vaut tout de même mieux que si elle était entièrement inconnue. C'est au point de vue abstrait du mathématicien, pour lequel les nombres sont considérés en dehors de leur rapport avec la réalité, qu'une erreur d'un millième de millimètre est chose aussi grave qu'une erreur d'un kilomètre : car si deux nombres a et b sont différents, que leur différence soit très grande ou très petite, on pourra arriver à des déductions également absurdes si l'on pose $a = b$ et si l'on transforme cette égalité par une série de calculs : c'est à ce point de vue que s'est placé Henri Poincaré dans ses spéculations sur la relativité de la connaissance. Mais quel que puisse être l'intérêt métaphysique de telles spéculations, elles ne doivent pas être mêlées à une question proprement scientifique ; dans une telle question, on ne doit pas confondre l'imprécision avec l'erreur totale. Toutes nos connaissances pratiques, qu'elles soient vulgaires ou scientifiques, sont imparfaites, mais cette imperfection n'est pas incompatible avec la certitude. Si je conviens de faire un signal à midi, un observateur attentif sera *certain* de le voir en guettant cinq minutes d'avance et cinq minutes après midi, si nos montres sont passables et ont été mises d'accord assez grossièrement dans la matinée. Avec des chronomètres de précision réglés avec soin, on pourrait réduire à quelques secondes la durée du guet ; dans tous les cas, la connaissance de

l'heure du signal projeté est un renseignement précieux, bien qu'il y ait des difficultés inextricables à définir cette heure avec une rigueur absolue. De même, ce serait un résultat positif de savoir, si l'on pouvait prétendre à une telle précision, que la taille moyenne d'un groupe humain est comprise entre $1^{m},653427$ et $1^{m},653429$; la légère incertitude provenant de l'impossibilité d'une définition rigoureuse du groupe n'enlèverait rien à la valeur absolue du résultat atteint.

56. — On conçoit maintenant en quel sens il peut être possible de rapporter à un étalon fixe les variations d'une espèce, cet étalon fixe étant déterminé par l'espèce variable elle-même. A chaque instant cette espèce se modifie ; certain individus naissent, d'autres deviennent adultes, d'autres meurent ou émigrent en dehors du cadre des observations.

Mais ces fluctuations, si elles rendent logiquement impossible la détermination rigoureuse de la valeur moyenne d'un caractère tel que la taille, sont cependant assez faibles pour que cette valeur moyenne puisse être connue avec une très grande précision. Il sera donc aisé de se rendre compte si cette valeur moyenne est la même ou non dans des groupes différents ou dans un même groupe à des époques différentes. Tel est le principe des méthodes par lesquelles on peut tenter l'étude précise des problèmes de l'évolution : elles sont basées sur l'étude statistique d'un grand nombre d'observations particulières. Chaque observation individuelle est, en réalité, dépourvue d'intérêt ; c'est un fait, mais ce n'est point un fait scientifique, de savoir que la taille de tel individu est

de $1^m,75$; ce qui est intéressant, c'est de savoir, sur 100.000 individus, combien il y en a dont la taille est de $1^m,75$ (ou, plus précisément, est plus voisine de $1^m,75$ que de $1^m,74$ ou de $1^m,76$).

Le caractère essentiel de la connaissance scientifique est de permettre la prévision; on pourrait proposer comme idéal à la biologie de déduire de la connaissance parfaite des parents la connaissance parfaite de l'enfant. Mais cet idéal est visiblement chimérique ; le simple fait que deux jumeaux peuvent être de sexes différents montre combien nous sommes loin de savoir seulement poser le problème. Ce qui est très remarquable, c'est que ce problème insoluble devienne relativement simple lorsqu'on porte simultanément l'attention sur un grand nombre d'individus. Aucun savant n'est en état de dire si un enfant qui naîtra dans quelques mois sera un garçon ou une fille, mais tous ceux qui s'occupent de démographie connaissent la proportion exacte entre le nombre des naissances masculines et des naissances féminines et sont scientifiquement certains que, l'an prochain, en France, il naîtra un peu plus de garçons que de filles. Cette même régularité se retrouve dans les phénomènes d'hérédité auxquels on a donné le nom de phénomènes mendéliens, du nom du moine autrichien Mendel, qui en a signalé les premiers exemples[1]. Elle se retrouve aussi dans

1. Voici en quoi consiste l'hérédité mendélienne; certains croisements entre variétés différentes donnent des produits (métis) qui ne se distinguent pas en apparence de l'une des variétés, ce que l'on exprime en disant que le caractère qui appartient à cette variété est dominant; c'est seulement en croisant ces métis entre eux qu'on constate qu'ils n'étaient pas de race pure. Supposons, pour un instant, que les croisements entre blancs et nègres satisfassent aux lois

les valeurs moyennes d'un grand nombre de mesures analogues, telles que la taille de tous les conscrits : nul ne peut prévoir quelle sera dans vingt ans la taille d'un enfant qui a actuellement trois ans; mais on peut prévoir quelle sera la taille moyenne d'un grand nombre de Français de vingt-trois ans.

La possibilité de prévision devient encore plus invraisemblable si l'on porte l'attention non sur les individus eux-mêmes, mais sur les germes innombrables susceptibles de devenir des êtres vivants; par exemple, sur les grains de pollen des champs de blé de France; leur nombre ne se chiffre pas seulement par milliards, mais par milliards de milliards. Etudier individuellement chacun d'eux serait une tâche surhumaine; ce qui est à la fois plus aisé et plus intéressant, c'est l'étude des propriétés globales de la récolte, et des modifications qu'apportent à ces propriétés les perfectionnements de la culture et l'introduction de nouvelles variétés de semences.

En résumé, certains problèmes sans intérêt et

de Mendel (ce qui n'est pas le cas); de mariages entre blancs et négresses ou entre nègres et blanches naîtraient alors des mulâtres en apparence identiques aux blancs. Mais si l'on croisait ces mulâtres entre eux, on obtiendrait invariablement, sur 1.000 produits, le quart (environ) soit 250 blancs purs. le quart (250) nègres purs et la moitié (500) mulâtres, c'est-à-dire en apparence blancs, mais que l'on reconnaîtrait être mulâtres à ce que les produits de leurs croisements mutuels suivraient la loi que nous venons d'indiquer, tandis que les blancs purs auraient des descendants exclusivement blancs. En fait, les choses ne se passent pas ainsi dans ce cas-là, mais il existe des espèces animales ou végétales pour lesquelles on peut faire une narration tout à fait analogue; il suffirait, par exemple, de remplacer, dans ce qui précède, les blancs par des tomates à chair rouge, les nègres par des tomates à chair jaune, et les mulâtres par des tomates à chair rouge métis (de race non pure). ne se distinguant des tomates à chair rouge proprement dites que par leur descendance.

pratiquement insolubles pour les individus conduisent à des recherches utiles et à des lois scientifiques lorsqu'on envisage simultanément un nombre très grand de cas, et que l'on porte l'attention sur les propriétés moyennes ou, comme on dit aussi, sur les propriétés statistiques. En matière d'hérédité, il n'est de lois scientifiques que des lois statistiques.

57. — On a donné le nom de *Biométrique* à l'ensemble des recherches dans lesquelles sont utilisées des ensembles de mesures effectuées sur des êtres vivants. Je ne ferai pas ici l'histoire de la biométrique, pour laquelle je renverrai à une substantielle étude de M. Vito Volterra[1]; j'ai déjà cité le nom de Quételet, qui fut un précurseur et ceux des Anglais Francis Galton et Karl Pearson.

Les problèmes que se pose la biométrique sont de natures diverses; indiquons deux des plus importants : l'homogénéité des groupes et les corrélations.

Supposons que l'on mesure la taille d'un grand nombre de Français adultes : si l'on convient de considérer la valeur moyenne des mesures comme la valeur exacte que devrait avoir la taille d'un Français, nous avons dit que les « erreurs », c'est-à-dire les différences positives ou négatives entre cette valeur théorique et la valeur réelle se répartissent précisément suivant la « loi de Gauss ». Nous avons convenu d'exprimer ce fait en disant que l'ensemble des

1. Vito Volterra. Sur l'application des mathématiques aux sciences biologiques et sociales (*Revue du Mois* du 10 janvier 1906, t. I, p. 1).

tailles des Français constitue une série *normale*. Biologiquement, ceci coïncide avec le fait que les Français constituent un groupe biologique suffisamment homogène.

Supposons que nous fassions la même expérience dans une ville dont les habitants appartiennent à deux races différentes, les uns étant blancs et les autres jaunes. Si l'on mesure tous les habitants adultes et que l'on note les mesures sans distinguer les races, on constate que la moyenne n'a plus la même signification et que le groupement autour de la moyenne ne suit pas de loi simple telle que celle de Gauss. Il y a superposition de deux phénomènes distincts, dont chacun suit la loi de Gauss, mais dont la somme ne suit pas cette loi ; les tailles des blancs se groupent autour de leur propre moyenne suivant la loi de Gauss et de même les tailles des jaunes autour de leur moyenne; la moyenne générale n'a aucune signification ; en particulier, il peut très bien arriver que cette moyenne générale ne soit pas la taille la plus fréquente. Ceci correspond au fait biologique que les races sont distinctes, et ce fait, s'il n'avait pas été connu, aurait pu être décelé par l'étude des moyennes.

Sans entrer dans des détails dont l'étude approfondie mériterait à elle seule tout un ouvrage, on conçoit comment l'étude expérimentale de faits analogues a pu conduire à la conclusion générale suivante : *le caractère biométrique intrinsèque des séries normales est la pureté de la race.* Je laisse de côté les commentaires nombreux dont aurait besoin un énoncé aussi général ; bien des idées auraient à être précisées, ce qui

est l'affaire du biologiste encore plus que du mathématicien; néanmoins, telle quelle est, cette théorie a prouvé sa valeur scientifique en se rendant utile par des applications pratiques. Je ne puis les mentionner toutes; je citerai toutefois les travaux très intéressants sur les orges de brasserie qui ont été inaugurés par M. Hjalmar Nilsson au laboratoire de Svalöf en Suède, et acclimatés en France par M. Blaringhem, dont les recherches personnelles ont été des plus intéressantes. On arrive, par l'application du caractère des races pures, à sélectionner et à cultiver dans des régions étendues des races d'orge extrêmement pures, dont la fixité est très appréciée par les brasseurs qui les emploient : on sait quels sont les avantages industriels de l'homogénéité de la matière première. On trouvera d'autres exemples dans le livre bien connu de M. Hugo de Vries *Espèces et Variétés*[1].

On voit combien est grande l'importance de la théorie biométrique de l'homogénéité; la théorie des corrélations n'est pas moins intéressante, car elle peut donner la clef des principaux problèmes de l'hérédité. On effectue diverses mesures sur plusieurs individus d'une même race, sur leurs parents et sur leurs collatéraux; quelles sont les corrélations de ces diverses mesures entre elles? Par exemple, connaissant la taille d'un individu, peut-on en conclure quelque chose relativement à la dimension de l'avant-bras de son frère? On voit que le concept de *cause* s'élargit ici singulièrement; on ne dira habituellement pas que la taille élevée d'un

1. Bibliothèque scientifique internationale (Félix Alcan).

individu est la *cause* de la longueur de l'avant-bras de son frère; cette extension de la notion de cause est parfaitement légitime au point de vue de la théorie des probabilités[1]. Je me bornerai, pour l'étude de cette théorie des corrélations, à renvoyer aux travaux de Karl Pearson et de ses élèves, parus notamment dans les *Proceedings of the Royal Society of London* et dans le journal spécial *Biometrika*.

58. — Nous renverrons au chapitre IX, consacré à l'étude de la valeur scientifique de la théorie des probabilités, la discussion des applications que l'on peut en faire à des questions de psychologie. Pour terminer ce qui est relatif aux sciences biologiques dans le sens le plus général du terme, je voudrais indiquer brièvement sous quelle forme mathématique on peut poser les problèmes de probabilité statistique.

Pour ne pas introduire d'inutiles complications verbales, nous admettrons que l'on considère des séries d'observations portant chacune sur le même nombre de cas possibles (par exemple sur une population constante); dans chaque série, on note le nombre des cas où s'est produit l'événement sur lequel porte l'observation; c'est ce

1. Dans une conception entièrement déterministe de l'univers, il n'y a pas de *cause* au sens habituel que l'on donne à ce terme : l'ensemble de l'univers doit, à tout instant, être regardé comme la cause de tous les événements passés, présents ou futurs; il n'est pas en effet possible de modifier un seul phénomène sans modifier tous les autres, car il faut pour cela concevoir *un autre* univers, celui dans lequel nous vivons ne pouvant, dans cette hypothèse, être conçu autre qu'il n'est. On doit donc, à ce point de vue entièrement déterministe, remplacer la notion de cause par celle de corrélation, qui s'applique également quel que soit l'ordre de succession des phénomènes étudiés et leurs relations directes apparentes.

qu'on appellera les *cas favorables*, cette appellation étant consacrée par l'usage (ce peuvent être les nombres des naissances, ou des décès, ou de personnes atteintes de telle maladie, etc.). L'observation fournit donc un tableau indiquant pour chaque série, le nombre des cas favorables ; on connaît d'autre part, le nombre des cas possibles. Le problème général de la statistique mathématique est le suivant :

Déterminer un système de tirages effectués dans des urnes de composition fixe, de telle manière que les résultats d'une série de tirages, interprétés à l'aide de coefficients fixes convenablement choisis, puissent avec une très grande vraisemblance conduire à un tableau identique au tableau des observations.

Nous donnerons au problème ainsi posé le nom de *problème du schéma des urnes*. Le type de la solution qu'il peut recevoir est le suivant : on considère trois urnes U_1, U_2, U_3, renfermant des boules identiques de forme, noires et blanches, la proportion des boules noires étant respectivement p_1, p_2, p_3 ; on fait dans l'urne U_1 un nombre d'extractions égale à α_1 (en remettant chaque fois la boule extraite), qui donnent n_1 boules noires, on extrait de même, en α_2 tirages, n_2 boules noires de U_2 et, en α_3 tirages, n_3 boules noires de U_3 ; le nombre des cas favorables est alors représenté par la somme

$$\lambda_1 n_1 + \lambda_2 n_2 + \lambda_3 n_3$$

les coefficients λ_1, λ_2, λ_3 ainsi que les nombres entiers α_1, α_2, α_3, et les probabilités p_1, p_2, p_3 définissent le schéma d'urnes considéré.

Nous devons nous contenter d'indiquer ici les diverses questions qui se posent à propos du problème général du schéma des urnes.

1° *Ce problème est-il possible?* Comme il s'agit seulement d'un problème d'approximation, il est clair qu'il est toujours possible en supposant le nombre des urnes suffisamment grand; mais une telle solution ne serait pas très satisfaisante; pratiquement, on doit chercher à le résoudre avec des urnes en nombre minimum; dans les cas où l'on peut y parvenir avec une seule urne, le problème considéré est dit *normal;* un cas intéressant est celui où deux urnes sont nécessaires et deux seulement.

2° *Le problème admet-il plusieurs solutions?* Certainement, si l'on suppose le nombre des urnes assez grand; mais, dans bien des cas, il n'y a qu'une seule *solution simple*, et c'est à celle-là qu'on s'attache particulièrement.

3° *Par quelle méthode obtenir les solutions?* Je me contenterai de renvoyer à ce sujet aux travaux de Karl Pearson, bien qu'il ne pose pas le problème sous la forme que nous lui avons donnée : la question me paraît d'ailleurs appeler de nouvelles recherches.

Pour terminer ces rapides indications, je tiens à signaler l'analogie entre la forme que nous avons donnée au problème du schéma des urnes et la forme sous laquelle se posent les problêmes de mécanique et de physique mathématique ; là, aussi, la première question est de construire un schéma mathématique présentant avec la réalité d'assez étroits rapports. Cette première question résolue, ou du moins partiellement résolue (car il y a souvent plusieurs solutions), le rôle du

mathématicien se borne à étudier les propriétés du schéma obtenu, ce qui est un problème de mathématiques pures; la comparaison des résultats obtenus ainsi avec l'expérience et le développement des théories que peut suggérer cette comparaison sont en dehors du domaine des mathématiques; car on doit, dans ces recherches théoriques, ne jamais perdre de vue les réalités et contrôler à chaque instant les idées nouvelles par l'observation et l'expérience. Mais le rôle des mathématiques, pour être limité, n'en est pas moins, dans bien des cas, fort important.

1. Pour l'application de la méthode du schéma des urnes a un problème biologique particulier, celui du sexe dans les accouchements doubles; voir mes *Eléments de la théorie des probabilités*, n° 62. On arrive à cette conclusion que tout se passe comme si, 28 fois sur 100 environ, le sexe de l'un des jumeaux était déterminé par le sexe de l'autre; tandis que 72 fois sur 100 environ, les sexes des deux jumeaux sont soumis aux mêmes lois de hasard que s'il s'agissait de deux naissances indépendantes l'une de l'autre. Il serait intéressant de se rendre compte si ces proportions sont celles mêmes dans lesquelles se partagent les naissances doubles entre naissances provenant d'un même œuf ou de deux œufs différents.



CHAPITRE VI

LES SCIENCES PHYSIQUES

59. — Il semble que les méthodes statistiques ne puissent pas s'appliquer aux sciences physiques de la même manière qu'aux sciences biologiques, puisque la notion d'individu fait défaut. Un litre d'air ne se distingue pas d'un autre litre d'air comme un grain de blé se distingue d'un autre grain de blé; deux dynamos de même construction fournissent des courants identiques. Les lois physiques ont précisément pour caractère d'exprimer des propriétés qui sont communes à une infinité d'êtres indiscernables entre eux. Cette manière de voir a été longtemps admise sans discussion. Ceux-là mêmes qui regardaient la matière comme formée d'un très grand nombre de parties distinctes, atomes ou molécules, ne pensaient pas qu'il y eût lieu de considérer ses

propriétés comme la résultante statistique des propriétés des atomes. Il semble que l'illustre physicien anglais Maxwell soit le premier à avoir introduit systématiquement les calculs statistiques dans l'étude des gaz [1]. Ce point de vue s'est montré très fécond et c'est à lui que l'on peut rattacher quelques-uns des progrès les plus importants qui sont en train de bouleverser la physique. Nous ne pouvons songer à donner en quelques pages un exposé de ces progrès, ni des formes diverses sous lesquelles intervient la méthôde statistique; mais il suffit d'un exemple très simple pour mettre en évidence les principes de cette méthode. Tout le monde sait qu'un gaz, enfermé dans une enceinte, exerce une pression sur les parois de cette enceinte, tend à la distendre si elle est élastique. Si nous gonflons un ballon avec de l'hydrogène ou avec du gaz d'éclairage, ce ballon ne se distend pas indéfiniment parce que la pression atmosphérique extérieure fait équilibre à la pression intérieure du gaz; mais si le ballon gagne des régions où la pression atmosphérique est plus faible, l'expansion nouvelle que pourra prendre le gaz augmentera le volume du ballon. Comment explique-t-on cette pression du gaz sur les parois, et les particularités qu'elle présente? Par exemple, elle est la même quelle que soit la direction de la paroi, c'est-à-dire que le plafond d'une chambre supporte la même pression que le plancher, tandis que si l'eau remplaçait l'air, il y aurait une pression sur le plancher, et non sur le pla-

1. Il convient de citer aussi les noms de Bernoulli, de Clausius et de Boltzmann.

fond, la pression de l'eau étant due à son poids.

On se représente le gaz comme formé d'un nombre extrêmement grand de molécules que l'on regardera, dans une première approximation, comme des sphères, des petites boules égales entre elles, qui sont dans une agitation perpétuelle, s'entrechoquent mutuellement et choquent les parois un grand nombre de fois par seconde.

Si l'on connaissait très exactement, à un instant donné, les positions de toutes ces molécules et leurs mouvements réciproques, le problème qui consisterait à prévoir les mouvements ultérieurs serait un problème de mécanique théoriquement très simple, bien que pratiquement insoluble en raison du très grand nombre des molécules. Il n'est même pas possible de songer à connaître individuellement chaque molécule ; la vie humaine est trop courte : elle ne dépasse guère un ou deux milliards de secondes ; un homme devrait penser à plusieurs milliards de molécules par seconde pour avoir pensé, à la fin de sa vie, à toutes les molécules d'une petite masse gazeuse[1].

Il est donc impossible de connaître les vitesses des molécules qui choquent à un instant donné une certaine portion de la paroi ; c'est cependant à ces chocs qu'est due la pression exercée par le gaz sur cette portion de paroi, et l'expérience nous apprend que cette pression reste la même lorsque les conditions physiques ne sont pas modifiées. C'est donc que les chocs, inconnus

1. Voir mon article sur *La philosophie mathématique et l'infini. Revue du Mois* du 10 août 1912, t. XIV, p. 218.

dans le détail ne sont cependant pas complètement inaccessibles à notre esprit : nous pouvons arriver à connaître avec précision le résultat global de tous les chocs qui se produisent pendant une fraction de seconde ; ce résultat dépend de la valeur moyenne de certaines quantités dont la valeur exacte pour chaque molécule nous est forcément inconnue. Le résultat qu'exprime la loi physique est donc, en définitive, un résultat statistique.

On a créé des méthodes de calcul permettant de prévoir ces résultats statistiques, moyennes de phénomènes trop nombreux pour que l'analyse individuelle en soit possible ; c'est par des lois statistiques que l'on tend à expliquer tous les phénomènes physiques et que l'on arrive à découvrir des propriétés ignorées. Toute loi statistique est d'ailleurs seulement une loi approchée ; mais lorsque les individus sont aussi nombreux que les molécules d'un gaz, l'erreur possible est beaucoup trop faible pour qu'aucun moyen humain d'investigation puisse la déceler ; tout se passe donc comme si la loi approchée était rigoureusement exacte.

60. — Examinons, par exemple, ce qui se passe lorsque l'on met en communication par une large ouverture deux récipients, A et B, de même volume, remplis de deux gaz différents, à la même pression et à la même température. D'après la loi d'Avogadro, chacun des récipients renferme le même nombre de molécules, nombre qui, pour les récipients de dimensions usuelles, est de l'ordre de grandeur de 10^{22}, c'est-à-dire de plusieurs milliers de milliards de

milliards[1]. L'expérience montre que les chocs irréguliers des molécules ont rapidement rendu le mélange des gaz pratiquement homogène, c'est-à-dire qu'il est expérimentalement impossible d'y déceler une hétérogénéité appréciable. De ce fait expérimental, on conclut, indépendamment de toute hypothèse théorique sur la nature des chocs, que, au bout du temps assez court nécessaire à la diffusion, chaque molécule de l'un quelconque des deux gaz a la même probabilité de se trouver dans le récipient A et dans le récipient B. Tout doit donc se passer comme si, pour chaque molécule[2], on tirait à pile ou face pour décider si elle doit se trouver, à un instant donné dans A ou dans B. Or nous savons que pour 10^{22} parties de pile ou face, l'unité décimale d'écart est 10^{11}, c'est-à-dire cent milliards ; il y a donc une chance sur 10 pour qu'à un instant donné le récipient A contienne 10^{11} molécules de plus que le récipient B ; ce nombre correspond à la fraction 10^{11} du nombre total des molécules, c'est-à-dire au cent millième de millionième ; telle est la valeur de l'hétérogénéité que l'on peut s'attendre à rencontrer dans un mélange de quelques litres de gaz à la pression et à la température habituelles. Il est manifeste qu'elle est complètement inaccessible à la mesure expérimentale. Ce qui est plus frappant encore, c'est la rapidité prodigieuse avec laquelle décroît la probabilité

1. Sur ces divers points, voir le livre de M. Jean Perrin, *les Atomes* (Nouvelle Collection Scientifique).

2. Bien entendu, on suppose ici que les molécules ne sont pas liées entre elles par groupes, de sorte que les probabilités de se trouver dans A sont indépendantes pour les diverses molécules. S'il y avait des groupes indissolubles, ce seraient ces groupes que l'on devrait nommer molécules.

d'une hétérogénéité plus forte. Nous savons que la probabilité d'un écart égal à 10 fois l'unité décimale est égale à 10^{-100}, c'est-à-dire à un nombre décimal qui occuperait au moins deux lignes de cet ouvrage, la virgule étant suivie de 99 zéros, suivis eux-mêmes du chiffre 1. Un tel écart correspond à une hétérogénéité de l'ordre du dix-millième de millionième ou du dix-milliardième. Si l'on voulait abaisser l'hétérogénéité au cent-millionième, il faudrait prévoir un écart égal à 1000 fois l'unité décimale; la probabilité serait $10^{-1000000}$, c'est-à-dire un nombre décimal comportant après la virgule un million de chiffres, dont le dernier seul serait le chiffre 1, tous les autres étant des zéros; pour écrire un tel nombre avec les caractères employés pour cet ouvrage, il faudrait toutes les pages d'un volume plus gros que celui-ci.

Enfin, pour que l'hétérogénéité s'abaisse jusqu'au cent-millième, la probabilité s'exprimerait par un nombre analogue au précédent, mais comportant un million de millions de zéros ; pour l'écrire, il faudrait plus d'un million de volumes tels que celui-ci. De tels nombres dépassent tellement ce que notre imagination peut concevoir, qu'il n'est pas inutile de chercher, par quelque comparaison, à donner une idée de la petitesse de ces probabilités.

61. — Admettons, pour fixer les idées, que le nombre des caractères employés dans l'écriture française, en y comprenant les signes de ponctuation, etc. soit égal [1] à 100 ; un livre de dimension

1. Ce nombre est inférieur à 100 dans l'écriture ordinaire, en tenant compte des majuscules et des chiffres ; les machines à écrire

moyenne renferme moins d'un million de caractères ; quelle est la probabilité pour que l'on obtienne ce livre entier en choisissant les caractères au hasard, en les tirant au sort ? Il est clair que la probabilité est $\frac{1}{100}$ pour que la première lettre choisie soit la première du livre, elle est aussi $\frac{1}{100}$ pour que la seconde lettre choisie soit la seconde du livre ; comme les deux événements sont indépendants, la probabilité pour qu'ils se produisent tous deux est

$$\frac{1}{100} \times \frac{1}{100} = \left(\frac{1}{100}\right)^2.$$

Le même raisonnement peut être répété pour la troisième lettre, la quatrième, etc. S'il y en a un million, la probabilité pour que le hasard les fournisse toutes exactement s'obtient en faisant le produit d'un million de facteurs égaux chacun à un centième ; elle est égale à

$$\left(\frac{1}{100}\right)^{1.000.000} = 10^{-2.000.000}.$$

Si au lieu d'un seul livre nous en considérions un million, nous devrions faire le produit d'un million de facteurs égaux au nombre que nous venons de trouver ; le résultat serait $10^{-2\,000\,000\,000\,000}$. C'est un nombre analogue à celui que nous avons trouvé pour la probabilité

usuelles en prévoient 84 ; il serait supérieur à 100 si l'on tenait compte des caractères variés usités en imprimerie : italiques, caractères gras, lettres grecques minuscules et majuscules, gothiques (encore usitées en Allemagne), etc. : mais on verra qu'il n'y aurait presque rien à changer à notre calcul, et rien à nos conclusions, si l'on remplaçait 100 par 1.000 ou même par 10.000.

d'une hétérogénéité égale à un cent-millième [1], pour une hétérogénéité de l'ordre du dix-millième, on aurait une probabilité dans laquelle l'exposant négatif serait 100 fois plus élevé, c'est-à-dire qui correspondrait à cent millions de volumes reproduits par le tirage au sort des caractères.

Les explications précédentes suffisent pour justifier la comparaison du « miracle des singes dactylographes », dont nous aurons à faire un fréquent usage, car elle est particulièrement frappante.

62. — Concevons qu'on ait dressé un million de singes à frapper au hasard sur les touches d'une machine à écrire et que, sous la surveillance de contremaîtres illettrés, ces singes dactylographes travaillent avec ardeur dix heures par jour avec un million de machines à écrire de types variés. Les contremaîtres illettrés rassembleraient les feuilles noircies et les relieraient en volumes. Et au bout d'un an, ces volumes se trouveraient renfermer la copie exacte des livres de toute nature et de toutes langues conservés dans les plus riches bibliothèques du monde. Telle est la probabilité pour qu'il se produise pendant un instant très court, dans le récipient A, un écart de l'ordre du cent-millième dans la composition du mélange gazeux. Supposer que cet écart ainsi produit subsistera pendant quelques secondes revient à admettre que, pendant plusieurs années,

1. Nous négligeons le facteur 2 qui se trouve dans l'exposant; pour en tenir compte il faudrait multiplier par $\sqrt{2}$ la valeur de l'hétérogénéité, c'est-à-dire parler de 1 divisé par 70.000 au lieu de 1 divisé par 100.000. Comme nous nous proposons simplement de donner des comparaisons qui fassent image, une telle modification est sans importance.

notre armée de singes dactylographes, travaillant toujours dans les mêmes conditions, fournira chaque jour la copie exacte de tous les imprimés, livres et journaux, qui paraîtront le jour correspondant de la semaine suivante sur toute la surface du globe et de toutes les paroles qui seront prononcées par tous les hommes en ce même jour. Il est plus simple de dire que ces écarts improbables sont purement impossibles.

63. — L'étude des vitesses des molécules d'un gaz est un peu plus compliquée : nous devrons nous contenter, en indiquant les résultats, de chercher à faire comprendre la marche par laquelle ils ont pu être obtenus. Dès le XVIII[e] siècle, Bernoulli proposa d'expliquer par l'agitation des molécules la pression exercée par un gaz sur une paroi. Comme cette pression est proportionnelle à la température absolue[1], on est conduit à admettre que la force vive des molécules est elle-même proportionnelle à la température. Connaissant le nombre et par suite les dimensions des molécules, un calcul facile permet, de la valeur de la pression, de déduire la force vive moyenne, et par suite la vitesse moyenne[2]. C'est ainsi que l'on trouve que les molécules d'azote et d'oxygène dont est formé l'air que nous respirons ont, aux températures usuelles, une vitesse moyenne de 4 à 500 mètres par seconde. Est-il possible d'admettre que toutes les molécules ont la même

1. C'est-à-dire à la température au-dessus du zéro absolu, qui correspond à — 273° C.

2. Dans un calcul précis, il faut prendre garde que la moyenne des carrés des vitesses n'est pas égale au carré de la moyenne des vitesses. La confusion est permise au degré de précision qui nous intéresse ici.

vitesse, les directions seules variant? Maxwell fut le premier à voir clairement que les lois du choc ne le permettent pas et sa *loi de répartition des vitesses* est l'une des plus belles et des plus fécondes applications des lois du hasard à la physique.

Maxwell obtint cette loi pour la première fois par un raisonnement très critiquable et dont le défaut fut signalé par plusieurs géomètres, notamment par Joseph Bertrand, qui est revenu à plusieurs reprises sur ces critiques. Ces critiques sont partiellement justifiées et nul ne regrette que l'on possède actuellement d'autres démonstrations de la loi de Maxwell que sa démonstration primitive. Il me semble cependant que les critiques n'ont pas suffisamment insisté sur le point qui est, à mon avis, le plus faible dans la démonstration de Maxwell, à savoir l'emploi de l'intégration d'une équation fonctionnelle ; il faudra toujours se méfier beaucoup des résultats ainsi obtenus, même si l'établissement de l'équation fonctionnelle ne donne lieu à aucune objection[1]. On ignore en effet, en général, dans quelle

1. Voici le raisonnement de Maxwell ; soit $\varphi(x)\,dx$ la probabilité pour que la projection de la vitesse sur Ox soit comprise entre x et $x + dx$; les probabilités analogues pour les projections sur Oy et Oz seront $\varphi(y)\,dy$ et $\varphi(z)\,dz$ puisque le gaz est isotrope. La probabilité pour que le vecteur vitesse issu de l'origine ait son extrémité dans le parallélépipède x, y, z ; $x + dx, y + dy, z + dz$ est donc

$$(1) \qquad \varphi(x)\,\varphi(y)\,\varphi(z)\,dx\,dy\,dz.$$

Mais cette probabilité doit, pour les mêmes raisons d'isotropie, être une fonction de la distance $\sqrt{x^2 + y^2 + z^2}$ de ce parallélépipède à l'origine ; on a donc

$$(2) \qquad \varphi(x)\,\varphi(y)\,\varphi(z) = f(x^2 + y^2 + z^2).$$

Telle est l'équation fonctionnelle qu'il s'agit d'intégrer ; on obtient

mesure une petite erreur dans une équation fonctionnelle peut entraîner des modifications considérables dans les résultats que l'on obtient en intégrant cette équation.

64. — Boltzmann a démontré la loi de Maxwell par la considération des chocs des molécules ; la démonstration est trop compliquée pour trouver place ici ; mais nous pouvons au moins nous rendre compte du fait que ces chocs ne permettraient pas aux molécules de conserver toutes la même

aisément en prenant la dérivée logarithmique par rapport aux 3 variables :

$$(3) \qquad \frac{\varphi'(x)}{x\,\varphi(x)} = \frac{\varphi'(y)}{y\,\varphi(y)} = \frac{\varphi'(z)}{z\,\varphi(z)} = \frac{2f'}{f}.$$

Chacun des membres de ces égalités dépendant d'une variable différente ne peut être qu'une constante qu'on désignera par $-2k^2$ (car l'hypothèse qu'elle serait positive conduit à des conséquences physiques absurdes) ; on trouve ainsi

$$(4) \quad \varphi(x) = e^{-k^2 x^2} \quad f(x^2 + y^2 + z^2) = e^{-k^2 (x^2 + y^2 + z^2)}.$$

C'est la loi de Maxwell. La critique de Bertrand porte sur l'équation (1), car l'application du principe des probabilités composées n'est légitime que si les événements sont indépendants. Or, on n'est pas assuré que la probabilité pour que la projection sur Oy soit comprise entre y et $y + dy$ est indépendante de la probabilité pour que la projection sur Ox soit comprise entre x et $x + dx$. En fait, une fois la formule de Maxwell démontrée par une autre méthode, on constate bien que ces probabilités sont indépendantes (contrairement à l'assertion de Bertrand qui affirme *a priori* qu'elles ne peuvent pas l'être). Il ne serait peut-être pas impossible de démontrer cette indépendance sans utiliser la formule de Maxwell et la démonstration que nous venons d'exposer deviendrait alors mathématiquement correcte. Elle laisserait cependant prise à la très grave objection qui peut être faite dans tous les cas où on emploie des équations fonctionnelles ; si les équations (3), au lieu d'être rigoureusement exactes, au sens mathématique du terme, ne sont qu'approchées, comme c'est forcément le cas pour des équations physiques, il n'est plus possible d'en tirer la conséquence de la constance absolue de chacun des membres et l'on ne peut rien en déduire de précis.

vitesse, à supposer que cette circonstance ait pu se produire à un instant donné. C'est une conséquence des lois élémentaires du choc : si deux billes de billard égales dont les vitesses sont égales en valeur absolue, mais dirigées à angle droit se choquent de telle manière qu'au moment du choc la ligne des centres ait la même direction que la vitesse de l'une des boules, celle-ci s'arrête tandis que l'autre se dirige suivant la bissectrice des deux vitesses avant le choc, avec une vitesse dont le carré est égal à la somme des carrés de ces deux vitesses. En somme, la vitesse de l'une des boules devient nulle, tandis que l'autre est multipliée par $\sqrt{2}$. Au contraire, si les vitesses étaient dirigées toutes deux suivant la ligne des centres, le choc modifierait les directions de ces vitesses, mais non leurs valeurs absolues. Entre ces deux cas extrêmes il peut se présenter tous les cas intermédiaires, dans lesquels la force vive totale peut se partager entre les deux billes d'une manière absolument quelconque. On voit donc que, si l'on supposait les molécules animées initialement de vitesses égales en valeur absolue (et de directions variées, bien entendu), cet état serait modifié dès que chaque molécule aurait subi un choc; un second choc amènerait de nouvelles modifications et, comme on est conduit à admettre que chaque molécule subit plusieurs milliards de chocs par seconde, on voit qu'au bout d'un temps inappréciable, l'irrégularité la plus complète se serait produite dans les valeurs des vitesses. Lorsque l'on cherche à imaginer le chaos d'un tel mouvement dans lequel chaque molécule subit des milliards de chocs chaque seconde on est conduit à admettre que la distribution des vitesses

est la même que si elle était produite par le hasard [1], avec cette seule restriction que la force vive totale reste invariante [2]. On est ainsi conduit à étudier un pur problème de théorie des probabilités. Représentons la vitesse de chaque molécule par un vecteur OM ayant son origine en un point fixe O et équipollent à cette vitesse ; *il s'agit de trouver la distribution la plus probable des points M dans l'espace, sachant que la somme des carrés de leurs distances à l'origine a une valeur donnée.*

C'est là un problème de probabilités continues qui ne présente aucune difficulté particulière ; sa solution conduit à la loi de Maxwell. Cette loi peut être mise sous la forme suivante, lorsque les points tels que M sont extrêmement nombreux : la *densité* de leur distribution dans l'espace est proportionnelle à $e^{-k^2r^2}$, r désignant la distance OM. Cette loi est précisément la même que la loi de distribution *dans le plan* des points atteints sur une cible par un très grand nombre de tireurs visant le point O. C'est dans le voisinage du point O que les points sont le plus denses, mais il ne faudrait pas en conclure que ce sont les vitesses voisines de zéro qui sont les plus fréquentes ; un calcul facile montre que les vitesses les plus fréquentes sont celles qui

1. Cette vue est confirmée par l'étude du mouvement dans l'espace à $3n$ dimensions (n étant le nombre des molécules). Voir, par exemple, mon *Introduction géométrique à quelques théories physiques.*

2. Il ne peut être question, dans de tels chocs, de perte de force vive transformée en chaleur, puisque la chaleur n'est pas autre chose, en théorie cinétique, que la force vive des molécules.

sont voisines de $\frac{1}{k}$[1]. Les molécules qui ont une vitesse 10 fois plus forte sont extrêmement rares, puisque la densité des points qui correspondent à de telles vitesses est e^{-100}.

Je n'insiste pas sur les calculs divers qui peuvent être effectués grâce à la loi de Maxwell; de la force vive totale des molécules, connue d'après l'étude globale du gaz, on peut déduire la valeur moyenne de la vitesse, la valeur du carré moyen, etc.

65. — Une question plus intéressante est celle des fluctuations. Nous avons déjà dit quelques mots des fluctuations de densité qui peuvent se produire dans un mélange gazeux (et aussi, bien entendu, dans un gaz simple); nous avons constaté que ces fluctuations sont inaccessibles à l'expérience pour des volumes notables à la température et à la pression ordinaires ; mais il n'en est pas nécessairement de même pour des volumes microscopiques ou des gaz raréfiés, ou encore pour des dissolutions très étendues (si l'on admet l'analogie entre les solides dissous et les gaz). Avant de préciser la valeur de ces fluctuations de densité, observons qu'il peut se produire aussi des fluctuations de température ; par définition, la température d'une masse gazeuse est proportionnelle à la force vive totale de ses molécules et par suite à la force vive moyenne d'une molécule, c'est-à-dire au carré moyen de la vitesse.

1. En effet, le nombre des vitesses comprises entre r et $r + dr$ s'obtient en multipliant la *densité* $e^{-k^2r^2}$ par le volume $4\pi r^2 dr$ de la couronne sphérique ; si l'on cherche le maximum du produit $4\pi r^2 e^{-k^2r^2}$, on voit immédiatement qu'il a lieu pour $r = \frac{1}{k}$.

On pourrait dire, à la rigueur, que la vitesse d'une molécule isolée est définie ainsi par sa vitesse, de sorte que dans toute masse gazeuse, il y a des molécules extrêmement froides et d'autres dont la température peut être 5 ou 6 fois plus élevée que la température moyenne de la masse (il s'agit, bien entendu, de températures absolues; 27° centigrades correspondent à 300° absolus; une température 5 à 6 fois plus élevée est donc de 1500° à 1800° absolus, c'est-à-dire environ de 1200° à 1500° centigrades); la considération de ces molécules chaudes a été invoquée pour expliquer certains phénomènes chimiques.

Une petite portion de la masse gazeuse aura une température généralement un peu différente de la température de la masse; nous savons que l'ordre de grandeur de la fluctuation dont la probabilité n'est pas nulle est proportionnel à l'inverse de la racine carrée du nombre des molécules.

C'est M. Smoluchowski qui a le premier soumis au calcul cette importante question des fluctuations, dans des conditions où la vérification expérimentale pouvait être tentée. C'est ainsi qu'il a rendu compte des phénomènes d'opalescence qui se produisent au voisinage du point critique.

Il faudrait citer aussi les recherches théoriques de M. Einstein sur le mouvement brownien et leur belle confirmation expérimentale par M. Jean Perrin; je renverrai au livre de M. Perrin sur *Les Atomes*.

66. — Gibbs a donné le nom de *mécanique statistique* aux recherches dans lesquelles inter-

viennent à la fois la mécanique et la théorie des probabilités, recherches dont le type classique est la théorie cinétique des gaz, due surtout à Maxwell et à Boltzmann. Au point de vue physique, le traité de Gibbs n'ajoute rien d'essentiel aux travaux de ses illustres devanciers ; mais au point de vue théorique et mathématique, il y a incontestablement de grands avantages à exposer la « mécanique statistique » d'une manière indépendante de toute hypothèse particulière sur la théorie cinétique.

L'expression « mécanique statistique[1] » a d'ailleurs deux sens ; on peut, comme nous l'avons fait dans les pages précédentes, considérer une masse gazeuse déterminée comme formée d'un très grand nombre de molécules et étudier statistiquement les propriétés des mouvements de ces molécules : c'est le point de vue de Maxwell et le premier point de vue de Boltzmann. Mais Boltzmann et, après lui, Gibbs, ont adopté une conception plus générale et parfois plus féconde ; c'est à cette conception que l'on se réfère le plus souvent lorsqu'on parle de mécanique statistique. Au lieu d'une masse gazeuse, considérons un nombre très grand de masses gazeuses identiques à la première, du moins en ce qui concerne les propriétés accessibles à l'expérience ; mais les mouvements des molécules sont différents dans ces diverses masses. Si nous considérons un nombre assez grand de telles masses pour que toutes les possibilités soient réalisées en ce qui concerne les mouvements des molécules, la mé-

1. Voir dans l'*Encyclopédie des sciences mathématiques* l'article de P. et T. Ehrenfest, exposé en français par Emile Borel.

canique statistique sera l'étude des propriétés qui sont les plus fréquentes parmi ces innombrables masses gazeuses [1].

On voit aisément que si l'on considère une masse renfermant 10^{24} molécules, les hypothèses différentes que l'on pourra faire sur son état sont au nombre de $A^{10^{24}}$, A étant le nombre d'hypothèses possibles pour une molécule. Rigoureusement parlant, dans l'hypothèse habituellement admise jusqu'ici de la continuité, A serait infini; en tout cas A est un nombre très grand et, en l'élevant à la puissance 10^{24}, on obtient un nombre incomparablement plus grand [2] que cet exposant 10^{24}; à savoir un nombre *dont le nombre de chiffres est égal à* 10^{24}. Pour un nombre d'épreuves aussi colossalement élevé, les fluctuations possibles par rapport aux résultats les plus probables sont absolument négligeables; ou, en d'autres termes, des fluctuations non négligeables doivent être regardées comme impossibles, tellement elles sont improbables. Elles sont aussi improbables que le miracle des singes dactylographes. Il n'y aurait d'ailleurs aucune impossibilité logique, comme l'a fait observer M. Ehrenfest, à aller encore plus loin et, qualifiant la mécanique statistique des molécules

1. Il est d'ailleurs plus correct, au lieu de parler de masses gazeuses, de parler simplement de « modèles-mathématiques » d'un gaz, ou de « modèles » d'un gaz; car on peut concevoir ces modèles abstraits comme aussi nombreux que l'on veut, tandis qu'il peut sembler étrange d'imaginer réalisées des masses d'hydrogène si nombreuses que leur masse totale dépasserait énormément la masse de l'univers visible.

2. On peut remarquer en passant que l'influence de l'exposant 10^{24} est prépondérante par rapport à la valeur de A; si l'on prend $A = 10^{10}$ au lieu de $A = 10$, cela revient à élever 10 à la puissance 10^{25} au lieu de 10^{24}.

comme du *premier ordre* et la mécanique statistique des modèles d'un gaz comme du *second ordre* à étudier des mécaniques statistiques des troisième, quatrième,...., $n^{\text{ième}}$ ordres. Mais cette vue théorique ne paraît avoir eu jusqu'ici aucune utilité effective.

67. — La mécanique statistique peut être interprétée comme l'étude des possibilités diverses qui peuvent être déduites de données partiellement indéterminées. Cette indétermination partielle des données s'impose nécessairement à l'esprit, du moment que l'on se place en face de la réalité : c'est seulement à un point de vue purement abstrait que l'on pourrait essayer de concevoir un problème mécanique dans lequel les valeurs initiales seraient connues avec une précision *absolue*. Or, il est aisé de se rendre compte que, dans la théorie cinétique des gaz, la multiplicité des chocs a pour conséquence d'augmenter avec une très grande rapidité et dans une proportion énorme, les plus légères indéterminations.

Plaçons-nous, pour fixer les idées, dans l'hypothèse où les molécules sont regardées comme parfaitement sphériques et imaginons une seule molécule mobile qui heurterait les autres molécules supposées absolument fixes. Cette hypothèse simplificative permet de rendre les raisonnements plus brefs et ne change rien d'essentiel dans les conséquences que nous avons en vue, comme on peut s'en assurer par l'analyse détaillée du cas où toutes les molécules sont mobiles.

Ce que l'on sait des dimensions des molécules, de leurs distances mutuelles et du nombre de

chocs éprouvés par chacune d'elles[1] permet de se faire l'image suivante du phénomène que nous étudions, examiné à un très fort grossissement : des globes sphériques fixes, d'un décimètre de diamètre, sont distribués irrégulièrement dans l'espace de manière que la distance de chacun d'eux au plus voisin soit de quelques mètres environ; un globe semblable aux précédents se déplace en se réfléchissant sur ces obstacles fixes, avec une vitesse[2] telle que le nombre des chocs par seconde est de l'ordre du milliard.

On voit immédiatement que tout se passe comme si le centre de la sphère mobile, réduite ainsi à un point matériel, se déplaçait en se réfléchissant sur des sphères concentriques aux sphères fixes, mais de rayons doubles, cette réflexion ayant lieu suivant les lois de la réflexion de la lumière. Notre problème est donc équivalent à l'étude de la réflexion d'un rayon lumineux sur ces nouvelles sphères fixes.

Imaginons maintenant qu'au lieu d'un rayon lumineux unique, géométriquement défini comme une pure ligne droite, nous ayons un pinceau infiniment délié : nous pouvons nous représenter ce pinceau délié sous la forme d'un cône de révolution dont l'angle au sommet est extrêmement petit (nous reviendrons tout à l'heure sur l'ordre de grandeur de cet angle). Il est évident qu'en se réfléchissant sur la surface convexe des sphères, le pinceau lumineux s'étalera ; on peut

1. Voir Jean Perrin, *les Atomes*.

2. Bien entendu, notre image grossie doit être regardée comme purement géométrique et cinématique ; au point de vue dynamique, on n'aurait pas le droit d'appliquer ainsi une telle transformation par similitude.

se rendre compte sans calcul de l'étalement produit par plusieurs réflexions successives. Il suffit en effet de se représenter des boules sphériques polies, d'un ou deux décimètres de diamètre, situées à quelques mètres de distance les unes des autres et d'observer la réflexion des rayons lumineux sur une ou plusieurs de ces boules. On constate tout d'abord, en portant son attention sur l'une d'elles, que tous les objets environnants y apparaissent réfléchis, mais notablement rapetissés ; l'image de la boule voisine en particulier a une dimension angulaire qui, dans nos hypothèses, est à peu près le dixième des dimensions angulaires apparentes des boules (la distance de notre œil aux boules étant de quelques mètres, comme la distance des boules entre elles). Si l'on peut distinguer dans la petite image de la boule voisine, les images de certains objets réfléchis dans cette boule et que l'on voit ainsi après deux réflexions successives, ces images se trouvent naturellement rapetissées dans la même proportion.

L'observation que nous imaginons ne permet guère d'aller plus loin ; car les réflexions sur plus de deux boules rapetissent tellement qu'aucun détail n'est plus perceptible ; le globe du soleil est seul assez lumineux pour qu'on puisse le distinguer comme un point, dont le diamètre apparent est trop petit pour être évalué. Mais cette simple observation rend intuitif le résultat auquel conduirait un calcul détaillé : dans les conditions imaginées, chaque réflexion diminue dans la proportion d'environ un dixième le diamètre apparent des objets ; ainsi, après une centaine de réflexions, le diamètre apparent du soleil nous

apparaîtrait comme de l'ordre de grandeur de 10^{-100} (degrés) ; ceci veut dire que les rayons partis de deux bords opposés du soleil arriveraient dans notre œil en faisant entre eux cet angle extrêmement petit. En vertu de ce que l'on appelle quelquefois le principe du retour inverse des rayons lumineux (principe qui se réduit ici simplement au fait que dans les lois de la réflexion le rayon réfléchi et le rayon incident interviennent symétriquement), deux rayons issus de notre œil et faisant entre eux cet angle extrêmement petit aboutiraient, aprés les mêmes réflexions subies en sens inverse, aux bords opposés du soleil, c'est-à-dire feraient entre eux un angle 10^{100} fois plus grand que leur angle initial. Après un milliard de réflexions, c'est-à-dire au bout d'une seconde, l'angle initial se trouverait multiplié par $10^{1000000000}$, c'est-à-dire par un nombre que notre imagination est absolument impuissante à concevoir. D'ailleurs, dès que les bords opposés du faisceau initial seront arrivés, après des réflexions successives, à faire un angle assez grand, les diverses portions de ce faisceau ne se réfléchiront plus sur les mêmes sphères. En revenant à l'image des molécules, la molécule mobile que nous considérons ne heurtera pas les mêmes molécules. Si donc on se propose d'écrire l'histoire d'une molécule dans les diverses hypothèses qui correspondent à des vitesses initiales intérieures au cône infiniment délié que nous considérons, on sera amené au bout d'une faible fraction de seconde à distinguer entre diverses parties du faisceau initial, parties dont les histoires seront entièrement différentes les unes des autres, du moment que les molé-

cules heurtées ne sont plus les mêmes. Au bout d'un temps encore plus bref une nouvelle subdivision sera nécessaire, et ainsi de suite, de sorte qu'au bout de quelques secondes, les subdivisions successives, ayant augmenté suivant une loi exponentielle seront plus nombreuses que les molécules du gaz. En partant des diverses vitesses initiales intérieures au cône infiniment délié, on aboutit ainsi, au bout de quelques secondes, à regarder comme également vraisemblable que la molécule considérée heurte une quelconque des autres molécules, ce choc pouvant de plus avoir lieu en un point absolument quelconque de la surface de la molécule ainsi choquée. Le nombre des combinaisons des molécules deux à deux étant égal simplement au carré du nombre des molécules, on voit aisément qu'au bout d'un temps double, toute combinaison imaginable de chocs deux à deux devient également probable, si l'on fait sur les vitesses de toutes les molécules les hypothèses que nous avons faites sur l'une d'elles[1].

Précisons maintenant un peu le caractère de ces hypothèses. Nous avons imaginé les vitesses initiales intérieures à un cône dont l'angle au sommet était extrêmement petit; il nous suffisait que cet angle devînt appréciable après multiplication par $10^{1000000000}$.

Cherchons à évaluer, en partant de la loi de Newton sur l'attraction universelle, la déviation que ferait subir à une molécule, dans l'intervalle

1. Le temps nécessaire serait bien plus long si l'on voulait réaliser un mode d'association *déterminé* de *toutes* les molécules deux à deux; car dans l'évaluation du nombre de telles associations, le nombre des molécules figure en exposant.

de deux chocs, le déplacement d'une masse extrêmement petite située à une très grande distance. Sous l'action de l'attraction terrestre un corps pesant tombe en une seconde de 5 mètres environ ; en 10^{-10} seconde, il tombera d'une quantité 10^{-20} fois plus petite ; si, au lieu du globe terrestre, nous considérons une petite sphère concentrique de même densité et dont les dimensions linéaires seraient 10^{17} fois plus faibles (la circonférence d'un grand cercle étant 4 dix-millionièmes de millimètre au lieu de 40 millions de mètres), la masse de cette sphère étant 10^{51} fois plus faible que celle de la terre, la déviation serait encore 10^{51} fois plus petite, soit $10^{20} \times 10^{51} = 10^{71}$ fois plus petite. Transportons maintenant cette sphère minuscule au delà des extrémités de l'univers visible, en un point d'où la lumière met des milliards d'années pour nous parvenir, au lieu qu'elle mettrait un cinquantième de seconde pour venir du centre de la terre; la distance étant environ 10^{18} fois plus grande, l'attraction devient 10^{36} fois plus faible; la déviation est donc en définitive $10^{71} \times 10^{36} = 10^{107}$ fois plus faible. Enfin, par un calcul analogue que j'omets, on aurait la valeur de la déviation qui correspondrait, non plus à l'action de la sphère minuscule placée à cette distance prodigieuse, mais au simple effet d'un déplacement d'un millionième de millimètre dans la position d'une telle sphère à une telle distance. Même en accumulant ainsi les hypothèses de nature à rendre l'action aussi faible que possible, on n'arrivera même pas à diviser la déviation par 10^{200}, c'est-à-dire que les variations produites par une telle action dans la trajectoire d'une molécule sont colossalement grandes par rapport

à ce qui nous a été nécessaire pour notre raisonnement. La représentation d'une masse gazeuse par un modèle unique, formé de molécules dont les positions et les vitesses à un instant donné sont rigoureusement déterminées, est donc une *pure fiction abstraite;* on ne peut se rapprocher de la réalité qu'en imaginant un faisceau de modèles, c'est-à-dire en attribuant aux données initiales une certaine indétermination. Si faible que soit cette indétermination, si faible aussi que l'on suppose l'indétermination des forces extérieures, l'effet des chocs *disperse* très rapidement les faisceaux de trajectoires supposés infiniment déliés et le problème du mouvement ultérieur des molécules devient, en très peu de secondes, très indéterminé, en ce sens qu'un nombre colossalement grand de possibilités différentes sont *a priori* également probables. Bien entendu, en un instant précis, *actuel*, une seule de ces possibilités est réalisée, mais l'indétermination renaît aussi considérable dès que l'on se pose le problème de l'état à une époque *future*, même très voisine. La seule forme sous laquelle le problème puisse être posé et résolu est donc la forme statistique : le très grand nombre des éventualités possibles peut-il être séparé en deux groupes très inégaux, toutes celles qui sont réunies dans le groupe le plus nombreux ayant certains caractères communs ? C'est ainsi que si l'on envisage toutes les éventualités possibles pour la succession d'un milliard de parties de pile ou face, on peut constituer un premier groupe avec les cas où l'écart est inférieur à 1.000.000, où par conséquent le nombre de parties gagnées est compris entre 499.000.000 et 501.000.000, et

un second groupe avec tous les autres cas. Le premier groupe étant extrêmement plus nombreux que le premier, il est *infiniment vraisemblable* que l'événement qui sera réalisé appartiendra à ce premier groupe, c'est-à-dire possédera ce caractère que le rapport du nombre des parties pile au nombre des parties face sera compris entre 0,996 et 1,004. Les conclusions auxquelles aboutit l'application des méthodes statistiques à l'étude des problèmes de la théorie cinétique sont de même nature que la conclusion précédente; d'ailleurs, en raison du très grand nombre des molécules, ils sont encore plus précis; quant au sens qu'il faut attribuer aux mots *infiniment vraisemblable*, nous ne pouvons que renvoyer à la comparaison du miracle des singes dactylographes.

68. — Lorsque l'on pose sous la forme précédente les problèmes de la mécanique statistique, l'objection dite *objection de Loschmidt* peut être aisément levée. Cette objection est la suivante. L'application de la théorie cinétique à l'étude des phénomènes thermodynamiques conduit à rendre compte de phénomènes irréversibles, tel que l'établissement de l'équilibre de température entre deux corps mis en contact. Or, la théorie cinétique utilise des phénomènes mécaniques qui sont tous réversibles, c'est-à-dire que les équations qui représentent ces phénomènes ne sont pas modifiées lorsque l'on change le signe du temps. Sous une forme plus concrète, si l'on conçoit qu'à un instant donné, on fasse rebrousser chemin à chaque molécule en lui imprimant une vitesse exactement opposée à sa vitesse

actuelle, tout se passera comme si, le mouvement ayant été cinématographié, on projetait le film à l'envers, en commençant par les portions les plus récentes. *Il n'est donc pas possible*, objecte Loschmidt, d'*expliquer par un tel mécanisme réversible des phénomènes irréversibles*. Cette objection tombe lorsque l'on s'est bien rendu compte du caractère nécessairement statistique des explications mécaniques; on ne cherche pas à déterminer l'allure rigoureusement définie des phénomènes mécaniques moléculaires, mais l'allure la plus probable parmi toutes les allures possibles; cette indétermination de l'avenir est le principe même de la mécanique statistique; *mais on ne saurait parler d'indétermination du passé*, et c'est pourquoi la distinction entre le passé et l'avenir, c'est-à-dire la dissymétrie du principe de Carnot en ce qui concerne le signe du temps, n'est pas contradictoire avec une explication mécanique des faits thermodynamiques.

Nous reviendrons plus loin (ch. X) sur la discussion de la théorie de l'irréversibilité.

69. — On a beaucoup étudié dans ces dernières années des phénomènes importants et nouveaux, dans lesquels la théorie des probabilités intervient pour ainsi dire à chaque instant : ce sont les phénomènes de radioactivité. Je ne puis m'étendre ici sur l'historique de la découverte ni sur le détail de ces phénomènes, qui ont pris rapidement une si grande place dans la physique[1]. Il est cependant nécessaire d'en pré-

1. Voir, pour l'historique, dans la *Revue du Mois* du 10 janvier 1913, la *Conférence Nobel* 1903 par Pierre Curie et la *Conférence Nobel* 1912 par Mme Pierre Curie. Plusieurs publications

ciser brièvement la nature. Le caractère essentiel de la radioactivité paraît être la décomposition *spontanée* de certains atomes. Cette décomposition se distingue très nettement de la dissociation chimique d'une molécule en plusieurs atomes, par plusieurs caractères dont le principal peut-être est son *invariance* à l'égard de tous les agents physiques[1]. En d'autres termes, en un temps donné, une substance radioactive déterminée se trouve perdre, en vertu du phénomène de la radioactivité, une proportion rigoureusement déterminée de son poids. Tout se passe donc comme si chaque atome radioactif avait à chaque instant la même probabilité de se briser pendant la seconde suivante, cette probabilité ne pouvant être modifiée ni par les agents physiques (température, pression, champ électrique ou magnétique), ni par le vieillissement spontané de l'atome lui-même. Si l'on admet ce point de vue comme une interprétation exactement adéquate des faits expérimentaux — et il semble bien qu'on ne puisse point ne pas l'admettre[2] — l'étude mathématique des phénomènes radioactifs est manifestement du domaine de la théorie des probabilités. Il est d'ailleurs possible, pour beaucoup de ces phénomènes, de ne faire intervenir que le résultat global observé, la

périodiques sont, dans divers pays, spécialement consacrées aux recherches scientifiques sur la radioactivité ; en particulier, en France, le journal mensuel *Le Radium*.

1. De plus, la radioactivité est jusqu'ici entièrement irréversible, c'est-à-dire qu'on ne connaît aucun moyen de recombiner les atomes qui résultent de la désintégration d'un atome radioactif. Mais ce caractère négatif est peut-être dû seulement à l'insuffisance de nos moyens d'action.

2. Voir : Mme Pierre Curie, *Traité de radioactivité* (Gauthier-Villars).

désintégration d'une proportion fixe dans un temps donné, mais pour d'autres phénomènes, la théorie des probabilités est forcément mise en jeu d'une manière explicite : c'est le cas pour l'étude des scintillations et le dénombrement direct des particules α émises par le radium et d'autres substances radioactives.

Ces particules α ne sont pas autre chose que des atomes d'hélium ; on est conduit à admettre que, en une seconde, un atome de radium sur cent milliards environ se décompose en un atome d'hélium et un atome de niton. L'atome d'hélium se trouve projeté avec une très grande vitesse et donne lieu à des phénomènes électriques qui permettent de déceler sa présence. Si l'on ne prend pas de précautions spéciales, les particules α ainsi émises sont trop nombreuses pour qu'on puisse les observer individuellement. Un gramme de radium renferme en effet plusieurs dizaines de milliards de fois cent milliards d'atomes, de sorte qu'à la proportion de un sur cent milliards, il s'en désintègre plusieurs dizaines de milliards par seconde [1]. Pour réduire ce nombre à quelques unités par seconde, ou même à une unité pendant quelques secondes, on songe immédiatement à deux moyens, qu'on utilisera simultanément. Tout d'abord, réduire la masse employée, ce qui est facile par l'emploi de solutions étendues de sels radioactifs; ensuite, réduire par le moyen d'écrans successifs l'ouverture angulaire dans laquelle on étudie l'émission. La loi du hasard s'applique en effet tout aussi bien à la direction dans laquelle est émise la

1. Environ 34 milliards d'après les observations les plus récentes.

particule α qu'à cette émission elle-même; des portions équivalentes de la surface d'une sphère ont des chances égales de recevoir ces projectiles; si l'on sépare ceux qui correspondent au millième ou au dix-millième de la surface de la sphère, on en aura donc 1.000 ou 10.000 fois moins. Or, la surface d'un cercle d'un centimètre de diamètre est la dix-millième partie de la surface d'une sphère de 25 centimètres de rayon. On conçoit donc sans peine qu'il ait été possible, même en opérant sur des corps plus radioactifs que le radium, d'arriver à déceler expérimentalement les émissions de particules α et à mesurer les intervalles de temps qui les séparent. La répartition de ces intervalles de temps autour de leur valeur moyenne est un problème de probabilités continues. On peut en effet, la durée de l'expérience étant faible par rapport à la durée nécessaire pour une diminution appréciable de la masse radioactive utilisée, considérer que la probabilité de l'émission est constante; l'espérance mathématique du joueur qui recevrait une somme fixe par particule émise, est donc proportionnelle au temps. Le problème de probabilités continues peut être posé sous la forme géométrique suivante : *Sur une droite indéfinie sont marqués au hasard un certain nombre de points, de telle manière qu'il y ait en moyenne* hx *points sur une longueur* x*; quelle est la probabilité pour que la distance d'un point marqué à celui qui est situé immédiatement à sa droite soit supérieure à une longueur donnée* y. Un calcul facile[1] montre que cette probabilité est

1. Voir E. Borel, *Introduction géométrique à quelques théories physiques.* Note V (Gauthier-Villars.)

e^{-hy}. Si l'on mesure un grand nombre de distances (c'est-à-dire d'intervalles de temps écoulés entre deux émissions successives), on peut vérifier l'accord entre ce résultat et l'expérience. Cette vérification a été faite d'une manière satisfaisante[1], rendant par suite très vraisemblable le point de départ du calcul. L'étude des scintillations conduit aussi à des résultats très satisfaisants.

70. — Il faudrait un ouvrage entier pour exposer avec détail les nouvelles applications de la théorie des probabilités à la physique, tellement elles ont été nombreuses dans ces dernières an-années. Bornons-nous à énumérer rapidement les recherches expérimentales et théoriques de M. Paul Langevin sur les gaz ionisés[2] et sur la perméabilité magnétique des gaz[3], les travaux expérimentaux de M. Jean Perrin sur le mouvement brownien[4], les études sur le rayonnement de MM. J. D. van der Waals junior[5], H. A. Lorentz[6]

1. Je citerai notamment une expérience très complète faite par Mme Curie, avec l'aide de ses préparateurs, et non encore publiée au moment où j'écris ces lignes. Cette expérience a porté sur 10.000 émissions et l'étude numérique, faite avec le plus grand soin par Mme Curie, concorde admirablement avec les prévisions théoriques. Cette concordance est la preuve expérimentale la plus complète de l'invariance de la radioactivité.

2. *Thèse*, Paris, 1902.

3. *Journal de physique*, 4 (1905), p. 678 et *Annales de chimie et de physique*, 8e série. t. IV, 1905, p. 70. C'est la très belle théorie du paramagnétisme de M. Paul Langevin, qui a conduit M. Pierre Weiss à la conception du *magnéton*, ou atome de magnétisme, conception qui rend compte de nombreux faits expérimentaux et a permis d'en prévoir d'autres. Voir Weiss. Les moments magnétiques des atomes et le magnéton dans l'ouvrage : *Les idées modernes sur la constitution de la matière* (Gauthier-Villars, 1913).

4. Voir Jean Perrin, *Les Atomes* (Paris, F. Alcan).

5. *Dissertation*, Amsterdam, 1900.

6. *Emission und Absorption der Metalle* (Proc. Amst. 1903, p. 666)

et Einstein[1], que l'on peut rattacher d'ailleurs à des recherches plus anciennes de lord Rayleigh[2], les études sur la dispersion diffuse de la lumière par les molécules[3], les travaux de MM. A. Cotton et H. Mouton sur la double réfraction magnétique et électrique dans les fluides[4], l'étude de l'allure thermodynamique des gaz ultra-raréfiés[5], c'est-à-dire tels que le chemin moyen parcouru par une molécule entre deux chocs est grand par rapport aux dimensions du vase, de sorte que les chocs contre les parois jouent un rôle prépondérant par rapport aux chocs des molécules entre elles, et bien d'autres questions dont l'exposition dépasserait le cadre de cet ouvrage. Il me paraît cependant nécessaire de donner au moins quelque idée de la manière dont se posent trois questions particulièrement importantes dans la physique moderne : l'équipartition de l'énergie, la définition générale de l'entropie comme logarithme de la probabilité, et la théorie des quanta.

71. — Le principe de l'équipartition de l'énergie consiste, sous sa forme la plus simple en ce que, dans un mélange de gaz, la force vive

1. *Annalen der Physik* (4), 19 (1906), § 2.

2. J.-W. Strutt (Lord Rayleigh). On the resultant of a large number of vibrations of the same pitch and of arbitrary phase (*Scientific papers*, 1871, n° 6, p. 76 (1880), n° 68, p. 491).

3. H.-A. Lorentz, Académie d'Amsterdam, 25 juin 1910 ; A. Einstein, *Annalen der Physik*, 34 (1911).

4. *Bulletin de la Société française de physique*, 1910. Voir aussi Paul Langevin dans le *Radium*, 1910, p. 249. et A. Corbino, *Physikalische Zeitschrift*, 1910, p. 756.

5. On trouvera un exposé très clair de cette dernière question dans la conférence de M. Louis Dunoyer publiée par la Société française de physique dans le volume déjà cité : *Idées modernes sur la constitution de la matière*.

moyenne (c'est-à-dire le demi-produit de la masse par le carré de la vitesse) est la même pour les molécules des diverses espèces; la vitesse moyenne diminue donc en raison inverse de la racine carrée du poids moléculaire. C'est ainsi que, les molécules d'oxygène étant 16 fois plus lourdes que les molécules d'hydrogène, leur vitesse moyenne, dans les mêmes conditions de température est 4 fois plus petite; à la température ordinaire, cette vitesse moyenne est de 1.700 mètres pour l'hydrogène et de 425 mètres pour l'oxygène.

Sans entrer dans le détail des calculs, qui sont assez longs et minutieux, on peut se rendre compte que les molécules dont la force vive dépasse notablement la moyenne ont, par rapport aux autres, une probabilité relativement plus grande de heurter des molécules qui se présentent sur leur chemin comme des obstacles et de tels chocs entraînent naturellement une perte de force vive au détriment de la molécule rapide et un gain au profit des molécules ainsi heurtées. Il n'est donc pas possible que certaines molécules privilégiées aient constamment une force vive supérieure à la moyenne; l'effet des chocs est de ramener vers la moyenne les molécules qui s'en seraient écartées momentanément. On peut donc présumer (bien entendu, ceci n'est en aucune manière un raisonnement rigoureux) que toutes les molécules auront une force vive tour à tour inférieure et supérieure à la moyenne et que, par suite, la force vive moyenne sera la même pour toutes.

L'équipartition de l'énergie se vérifie expérimentalement, comme l'a montré M. Jean Perrin,

pour les particules microscopiques en suspension dans les liquides et qui sont agitées de mouvements irréguliers auxquels on a donné le nom de mouvement brownien. Nous venons de dire qu'à la température ordinaire la vitesse d'une molécule d'hydrogène est 1.700 mètres par seconde; cette vitesse est donc égale à 170.000 en unités C. G. S.; d'autre part, la masse de la molécule d'hydrogène est environ 3.10^{-24}; l'énergie cinétique s'obtient en multipliant la moitié de la masse par le carré de la vitesse, ce qui donne environ 4.10^{-14} dynes. Si l'on observe, comme l'a fait M. Perrin, des grains dont la masse est de l'ordre de grandeur de 10^{-14} (ce qui correspond à un diamètre de quelques dix-millièmes de millimètre), la vitesse correspondant à cette énergie cinétique sera de l'ordre de grandeur de l'unité, c'est-à-dire du centimètre par seconde.

En fait, on ne mesure pas la vitesse, mais le déplacement en un temps donné, déplacement qui est la résultante d'un grand nombre de déplacements irréguliers; la théorie des probabilités permet, comme l'a montré M. Einstein, de déduire la valeur de la vitesse moyenne de la valeur du déplacement moyen ainsi mesuré [1].

Nous n'avons parlé jusqu'ici que de l'énergie qui correspond au déplacement d'ensemble des molécules, ou *énergie de translation;* c'est à cette énergie que le principe de l'équipartition a été d'abord appliqué et c'est pour elle qu'il soulève le moins de difficultés. Mais on peut donner de ce principe un énoncé beaucoup plus général,

1. Voir l'ouvrage déjà cité de M. Jean Perrin.

qui conduit dans certains cas à des conséquences incontestablement exactes, mais qui soulève dans d'autres cas des difficultés très sérieuses. Cet énoncé repose sur la notion des degrés de liberté d'un système, notion fort claire pour les systèmes mécaniques simples, mais assez confuse pour les systèmes continus tels que certains physiciens imaginent l'éther. Une sphère supposée parfaitement polie ne peut en aucun cas être amenée, par une action mécanique extérieure, à tourner sur elle-même; ses déplacements se confondent donc avec les déplacements de son centre de gravité; s'il peut se déplacer librement dans l'espace à trois dimensions, on dira que la sphère possède trois degrés de liberté; elle n'en posséderait que deux si son centre était assujetti à rester dans un plan donné ou une surface fixe. Envisageons maintenant une molécule dont la forme serait celle d'un ellipsoïde à trois axes inégaux; les chocs auront pour résultat de faire tourner cet ellipsoïde et, pour fixer sa position, il faudra, outre les trois coordonnées du centre, connaître les directions des axes, ce qui exige trois nouvelles données numériques; l'ellipsoïde possède donc six degrés de liberté; ce nombre se réduirait à cinq si l'ellipsoïde était de révolution; il ne peut pas se réduire à quatre, car un ellipsoïde qui serait de révolution autour de deux de ses axes serait aussi de révolution autour du troisième et se réduirait à une sphère[1]. L'énoncé

1. C'est là une propriété dont l'origine est la structure du groupe des symétries de l'espace. Il est fort remarquable que cette propriété purement analytique et géométrique de ce groupe se retrouve dans les propriétés physiques dont nous allons parler dans un instant (chaleurs spécifiques).

général du principe de l'équipartition est le suivant : *l'énergie cinétique moyenne est la même pour chaque degré de liberté.* Une des confirmations expérimentales les plus remarquables de ce principe est fournie par les valeurs des chaleurs spécifiques des gaz et des solides à la température ordinaire; cette chaleur spécifique, rapportée à la molécule, prend des valeurs qui sont très sensiblement les mêmes pour des groupes nombreux de corps, ces valeurs constantes étant elles-mêmes proportionnelles aux nombres 3, 5 et 6, qui correspondent aux trois cas de symétrie que nous avons envisagés : sphère, ellipsoïde de révolution, ellipsoïde à trois axes inégaux. Mais l'étude des chaleurs spécifiques aux basses températures donne lieu à de grosses difficultés dont nous dirons quelques mots tout à l'heure à l'occasion de la théorie des quanta.

Il n'est pas possible d'entrer ici dans le détail des calculs relatifs au principe général de l'équipartition; indiquons cependant que la définition précise de l'énergie moyenne *relative à un degré de liberté donné*, exige nécessairement la décomposition de l'énergie totale en une somme de carrés, relatifs chacun à un degré de liberté, sans termes rectangles[1].

Une autre difficulté est plus sérieuse encore et ne semble pouvoir être résolue que par des hypothèses de discontinuité analogues à celles de

1. Dans certains calculs, on admet que les termes rectangles disparaissent dans les moyennes, par raison de symétrie. Ces calculs me paraissent devoir être examinés de très près, en raison du nombre très grand des termes rectangles par rapport aux termes carrés ; chacun d'eux peut donc être très petit et leur ensemble cependant appréciable.

la théorie des quanta. Il ne semble pas possible d'admettre qu'une molécule ou qu'un atome est *rigoureusement* sphérique et ne peut par suite acquérir *aucune* énergie de rotation. Or, si faible que soit la dissymétrie par rapport à la sphère, la rotation cesse d'être *impossible :* comme les occasions (chocs) sont innombrables, cette rotation se produira et l'équipartition jouera. On peut s'en rendre compte sans calcul en observant que les mêmes raisons qui rendent difficile l'acquisition de l'énergie de rotation en rendent difficile aussi la déperdition une fois qu'elle est acquise. Bien d'autres difficultés peuvent être soulevées à propos de l'équipartition ; elles sont loin d'être toutes résolues [1] ; ces difficultés ne doivent tout de même pas nous faire oublier les cas dans lesquels le principe de l'équipartition repose sur des bases solides ; ce principe peut alors être regardé comme traduisant une propriété géométrique des surfaces du second degré (ellipsoïdes) dans l'espace à un très grand nombre de dimensions, surfaces dont l'étude fait partie de la théorie des probabilités [2].

72. — Boltzmann et Gibbs ont montré que, dans le cas d'un gaz, la fonction à laquelle Clausius a donné le nom d'entropie se confond à un facteur près avec le logarithme de la probabilité d'un état déterminé. La forme donnée par Clausius au principe de Carnot, d'après laquelle l'entropie d'un système isolé ne décroît jamais, se

1. Voir les Rapports et Discussions de Bruxelles, 1911, publiés par Paul Langevin et Maurice de Broglie (Gauthier-Villars).

2. Voir Emile Borel, *Introduction géométrique à quelques théories physiques* (Gauthier-Villars).

traduit donc par cette proposition qu'un système isolé ne passe jamais d'un état plus probable à un état moins probable; il évolue spontanément vers les états les plus probables.

L'entropie d'un système formé par la réunion de deux autres est d'ailleurs égale à la somme des entropies, tandis que la probabilité d'un état déterminé pour le système total est égale au produit des probabilités des états correspondants pour les systèmes partiels; ces relations concordent avec la propriété fondamentale des logarithmes : le logarithme d'un produit est la somme des logarithmes des facteurs. On a été ainsi conduit à étendre aux cas les plus généraux la relation déduite de la théorie cinétique pour les gaz :

Entropie $=$ $h\times$ *logarithme de la probabilité.*

Il est d'ailleurs possible de considérer cette relation fondamentale comme la définition même de la probabilité, en considérant l'entropie comme définie (à une constante additive près) par la thermodynamique classique; on peut aussi, inversement, définir l'entropie en partant de la probabilité. Dans un cas comme dans l'autre, il se présente des difficultés assez sérieuses dans la définition même de la probabilité d'un état déterminé pour un système dépendant de paramètres qui varient d'une manière continue. Certaines de ces difficultés, mais non toutes, sont évitées au moyen de l'hypothèse des quanta de Planck, dont nous allons maintenant dire un mot.

73. — L'hypothèse des quanta a été imaginée pour éviter les conséquences contraires à l'expé-

rience auxquelles conduit l'application du principe de l'équipartition de l'énergie à la théorie du rayonnement. On sait que dans une enceinte close le rayonnement a pour effet d'établir un équilibre de température entre tous les corps renfermés dans l'enceinte; l'équilibre s'établit séparément pour les diverses radiations de longueurs d'onde différentes; c'est ainsi que l'on est conduit à définir ce que l'on appelle : *la composition du rayonnement noir à une température déterminée.* On se représente, d'autre part, l'émission par rayonnement comme due à des *résonateurs* dont chacun a une période (et par suite une fréquence) déterminée.

La composition du rayonnement noir à une température déterminée, équivaut donc à la répartition de l'énergie totale du rayonnement entre les résonateurs de diverses périodes. Si cette répartition se faisait suivant les lois de la mécanique statistique ordinaire, elle devrait satisfaire au principe de l'équipartition. Or, ceci entraînerait que l'intensité du rayonnement pour les diverses longueurs d'onde devrait varier proportionnellement à la température absolue; par suite, une plaque d'argent polie, qui est très brillante dans l'obscurité lorsqu'on la porte à 1.227° C. (ou 1.500° absolus) devrait être 5 fois moins brillante, c'est-à-dire briller encore d'une manière très appréciable à 27°C. (ou 300° absolus). Cette conclusion est en contradiction flagrante avec l'expérience immédiate. Il est donc nécessaire d'admettre que les résonateurs dont la période correspond aux radiations visibles émettent, à la température ordinaire, beaucoup moins d'énergie que ne l'exigerait le principe de l'équiparti-

tion. L'hypothèse par laquelle M. Planck a expliqué ce fait expérimental est au premier abord étrange ; elle s'est montrée cependant très féconde depuis dix ans et, malgré des difficultés nombreuses restant à résoudre, malgré les modifications de détail, parfois contradictoires entre elles, qu'il a fallu lui faire subir pour l'adapter à diverses théories, il semble bien incontestable aujourd'hui que cette hypothèse n'est pas une pure abstraction théorique, mais que la *discontinuité,* qui en est l'essentiel, subsistera dans les transformations dont ses plus fervents adeptes sont d'accord pour reconnaître la nécessité.

L'hypothèse essentielle de la théorie des quanta, c'est que, dans certaines circonstances tout au moins (émission), l'énergie d'un résonateur ne peut pas varier d'une manière continue, mais seulement par *quanta* discontinus, multiples entiers d'un quantum, qui est lui-même égal au produit de la fréquence par une constante h (constante universelle de Planck). Plus la fréquence μ est grande, c'est-à-dire plus les périodes sont courtes, plus le produit $h\mu$ est élevé ; on peut ainsi se rendre compte comment les résonateurs de grande fréquence, correspondant aux vibrations lumineuses, peuvent ne pas entrer en jeu à la température ordinaire, bien que le rayonnement calorifique soit très appréciable.

Au point de vue de la théorie des probabilités, la théorie des quanta introduit aussi une discontinuité et, dans l'étude de l'équipartition, remplace un problème de probabilités continues (qui conduisait à des conséquences contraires à l'expérience) par un problème de probabilités disconti-

nues. Ce problème est celui de la distribution la plus probable de l'énergie entre les divers résonateurs, cette distribution devant être faite de telle manière qu'il soit attribué à chaque résonateur un multiple entier du quantum défini par sa période (ou bien du moins en regardant comme identiques au point de vue de la probabilité des distributions dans lesquelles l'énergie attribuée à un résonateur est comprise entre un multiple entier d'un quantum et le multiple entier immédiatement supérieur).

On peut observer, d'ailleurs, que si l'on n'admet pas une discontinuité ou une hypothèse équivalente[1], la définition de la probabilité d'un état conduit à des difficultés analytiquement inextricables ; les seules répartitions dont la probabilité ne soit pas nulle sont les répartitions uniformes ou quasi-uniformes. L'introduction systématique de la discontinuité ne supprime pas toutes les difficultés dans la définition générale de la probabilité d'un état, mais elle les atténue ; on pourrait concevoir qu'en systématisant l'emploi des variables discontinues dans la définition des grandeurs physiques (quanta de rotation, etc.) les phénomènes qui peuvent se produire dans un espace limité et clos soient considérés comme dépendant d'un nombre fini de paramètres pouvant prendre chacun seulement un nombre fini de valeurs, le nombre total des possibilités étant ainsi fini, bien qu'à vrai dire extrêmement grand. Une telle conception ne s'accorde guère avec les idées de continuité qui se sont montrées si

1. Sur certaines hypothèses continues équivalentes aux hypothèses discontinues, voir E. Borel, *Comptes rendus de l'Académie des Sciences*, 5 janvier 1914.

fécondes, au XVIII[e] et au XIX[e] siècle, dans l'étude des phénomènes naturels et grâce auxquelles tant de conquêtes scientifiques et industrielles ont été réalisées. Il paraît bien difficile d'abandonner complètement ces conceptions anciennes et il n'est pas aisé de se rendre compte dans quelle mesure elles sont conciliables avec les idées nouvelles ; ce qui paraît incontestable, c'est le rôle important que ne cessera plus de jouer la théorie des probabilités dans les théories physiques.



CHAPITRE VII

LES SCIENCES MATHÉMATIQUES

74. — Les applications de la théorie des probabilités à l'astronomie peuvent être classées soit parmi les applications aux sciences physiques, en raison de leur analogie avec la théorie cinétique des gaz, soit parmi les applications aux sciences mathématiques, car la mécanique céleste, depuis Newton et Laplace, doit être regardée comme faisant partie des sciences mathématiques, puisqu'elle déduit d'un très petit nombre de principes des résultats précis et rigoureux que l'observation n'a plus qu'à vérifier.

Les recherches astronomiques et cosmogoniques dans lesquelles intervient la théorie des probabilités, sont trop spéciales pour qu'il soit possible de les exposer ici en détail et d'autre part n'ont pas conduit à des résultats assez définitifs pour qu'on puisse se contenter d'en résumer les conclusions. Je dois donc me borner à en signaler l'existence[1].

1. Parmi les travaux les plus importants sur ces questions, il faut tout d'abord signaler ceux de Lord Kelvin (voir ses œuvres

75. — Il est au premier abord paradoxal de penser que la théorie du hasard peut s'appliquer aux sciences mathématiques pures, puisque, semble-t-il, c'est là un domaine où tout est défini avec précision, où les conclusions sont absolument rigoureuses et où par conséquent il n'y a pas de place pour le hasard et pour la probabilité. Je voudrais essayer de faire comprendre comment on a pu être amené à introduire, dans des questions de mathématiques pures, le langage et les méthodes de la théorie des probabilités; je serai obligé, pour y parvenir, de demander à ceux des lecteurs qui ne connaissent des mathématiques que leurs éléments, de vouloir bien me faire confiance et admettre sans démonstration les résultats dont la justification est exposée dans les traités spéciaux et ne peut trouver place ici [1].

76. — Les questions de probabilité les plus simples que l'on puisse se poser en mathématiques pures sont les questions relatives aux nombres réels, car l'ensemble de ces nombres est l'un des ensembles infinis que nous connaissons le mieux, ou du moins que nous croyons connaître le mieux, car dès qu'on en approfondit la nature on se heurte bien vite, comme nous le verrons tout à l'heure, à des difficultés imprévues. Cependant, il n'est pas douteux que toute personne possédant les éléments des mathématiques com-

complètes). Tout récemment M. Charlier, astronome à l'Observatoire de Lund (Suède) a publié une série de notes et mémoires du plus grand intérêt sur la statistique stellaire.

1. Voir, par exemple, Emile Borel, *Leçons sur la théorie des fonctions* (Gauthier-Villars).

prend de quoi il s'agit lorsqu'on lui parle des nombres compris entre 0 et 1, c'est-à-dire de l'ensemble des fractions décimales illimitées telles que:

0,43752.....

Parmi ces nombres, on distingue ceux qui sont commensurables (sous-entendu: *avec l'unité*), c'est-à-dire qui sont égaux à une fraction telle que $\frac{1}{2}$ ou $\frac{2}{3}$, et ceux qui sont incommensurables, c'est-à-dire qui ne sont égaux rigoureusement à aucune fraction, comme c'est le cas par exemple pour la racine carrée de 2; l'incommensurabilité de la diagonale du carré au côté de ce carré était déjà connue des géomètres grecs. Lorsqu'on écrit les nombres sous forme de fraction décimale, les nombres commensurables donnent lieu à des fractions limitées telles que 0,5 ou à des fractions périodiques telles que 0,166..... qui est égal à $\frac{1}{6}$. La fraction limitée peut d'ailleurs être regardée comme un cas particulier de la fraction périodique: le cas où la période serait formée ou exclusivement de zéros ou exclusivement de 9 (car 0,4999..... est équivalent à 0,5). Il est à peu près évident que la probabilité pour qu'une fraction décimale illimitée soit périodique, est égale à zéro [1], car cette probabilité est plus petite que tout nombre assignable. Si l'on imagine en effet, que l'on écrive les chiffres décimaux succes-

1. Pour démontrer ce point en toute rigueur, il faut utiliser la théorie des probabilités dénombrables. Voir E. Borel, Les probabilités dénombrables et leurs applications arithmétiques (*Rendiconti del Circolo matematico di Palermo*, t. XXVII, 1909), et Un problème de probabilités relatif aux fractions continues (*Mathematische Annalen*, t. 72, 1912).

sifs à partir de la virgule, il faut, pour que la fraction décimale soit périodique, qu'à partir d'un certain moment [1] tous les chiffres se reproduisent dans un ordre déterminé ; il faut donc qu'il se produise *une infinité* d'événements favorables ; en vertu du principe des probabilités composées, on doit faire le produit d'une infinité de fractions plus petites que l'unité [2] ; le résultat est donc aussi petit que l'on veut, c'est-à-dire zéro. On pourrait exprimer le même fait en disant que la probabilité est *infiniment petite*, les mots *infiniment petit* n'étant pas pris dans le sens de très petit, mais dans un sens voisin de celui qu'on leur donne en analyse mathématique : quantité variable qui tend vers zéro. Il peut sembler singulier en effet, de parler de probabilité zéro pour un événement qui n'est pas impossible, *car il y a* des nombres commensurables ; il n'est cependant pas paradoxal de le faire, si on observe que l'ensemble des nombres est infini et que le produit de l'infini par zéro doit être regardé non comme nul, mais comme indéterminé. C'est ainsi que la probabilité pour qu'un nombre choisi au hasard soit précisément un nombre donné, est évidemment nulle [3].

1. Le raisonnement rigoureux et complet exigerait que l'on tienne compte du fait que la période peut être aussi longue que l'on veut ; nous renverrons au mémoire qui vient d'être cité.

2. Et même plus petites qu'une fraction fixe inférieure à 1.

3. Bien entendu, ceci suppose que l'on s'entende sur le sens des mots *choisi au hasard* ; nous admettons qu'ils signifient que la probabilité est la même pour tous les nombres ; sa valeur ne peut donc être que zéro, car si elle était positive, si petite qu'elle fût, on obtiendrait pour le choix d'un nombre suffisamment grand de nombres une probabilité supérieure à l'unité, ce qui est absurde. Il n'en est pas de même si l'on considère que le nombre inconnu dont on évalue la probabilité est la solution d'un problème d'ana-

77. — La conclusion à laquelle nous avons été conduits par l'étude arithmétique des nombres décimaux se trouve justifiée aussi par la représentation géométrique des nombres sur une droite fixe ou *axe* par le point dont le nombre est l'abscisse.

Le nombre x est ainsi représenté par un point M de l'axe tel que OM soit égal à x, O étant l'origine des abscisses [1]; les nombres x compris entre 0 et 1 sont ainsi représentés par tous les points d'un segment OA dont la longueur est égale à l'unité, le point O correspondant à 0 et le point A à 1. Il est clair que toute portion, si petite qu'elle soit, du segment OA, renferme une infinité de nombres rationnels (c'est-à-dire de points représentatifs de nombres rationnels) ; car si une portion MN est supérieure, pour fixer les idées, à un millionième, il suffira de considérer les intervalles successifs d'étendue égale à un dix-millionième, à savoir 0 à 0,0000001, puis 0,0000001 à 0,0000002, etc. jusqu'à 0,9999999 à 1, pour qu'il soit certain que plusieurs de ces intervalles seront entièrement intérieurs à MN (sinon la longueur de MN serait inférieure à la somme de deux intervalles consécutifs, c'est-à-dire à deux dix-millioniėmes) ; si l'on suppose que l'intervalle 0,345 6778 à 0,3456779 soit intérieur à MN, tous les nombres rationnels dont les sept premiers chiffres décimaux

lyse, même compliqué. En ce cas, il est à peu près impossible de définir correctement la probabilité, car on pressent, sans qu'il soit aisé de justifier avec précision cette vue *a priori*, que les nombres simples tels que les nombres rationnels, ou les irrationnelles élémentaires, ou les nombres transcendants usuels tels que e ou π sont *plus probables* qu'un nombre dont les chiffres décimaux seraient régis par une loi arithmétique très compliquée.

1. Je laisse de côté les nombres négatifs.

sont 0,3456778 sont sur MN ; il y en a évidemment une infinité. Le raisonnement serait le même si MN, au lieu d'être supposé supérieur à un millionième, était supposé supérieur à un milliardième de milliardième, ou à toute autre fraction si petite soit-elle.

On exprime cette propriété en disant que les nombres rationnels sont *denses* dans tout l'intervalle OA. Comment ce fait est-il compatible avec celui que la probabilité est nulle pour qu'un nombre choisi au hasard soit rationnel ? Pour s'en rendre compte, il est nécessaire de préciser une notion fort importante, celle de la *mesure* d'un ensemble linéaire, c'est-à-dire d'un ensemble formé de points situés sur un segment tel que OA. Par définition, la *mesure* d'un tel segment sera sa longueur, que nous supposons égale à l'unité ; si l'on a un segment MN situé sur OA et dont la longueur est 0,3 par exemple, on dira de même que la mesure de l'ensemble des points du segment MN est égale à 0,3. Cette définition s'étend sans difficulté au cas où l'on considère sur OA un certain nombre limité de segments MN, M′N′, M″N″ n'empiétant pas les uns sur les autres ; la *mesure* de l'ensemble formé par les points qui appartiennent à ces segments est égale à la somme de leurs mesures, c'est-à-dire de leurs longueurs, et cette somme est forcément inférieure à 1, mesure de OA, du moment que les segments sont tous sur OA et n'empiètent pas les uns sur les autres. Avec ces définitions, il est clair que la probabilité pour qu'un point P *assujetti à se trouver sur* OA, soit sur MN est égale à la mesure de MN ; d'une manière générale, la probabilité pour que le point P appartienne à un

ensemble formé des segments MN, M′N′, M″N″ est égale à la mesure de cet ensemble ; la probabilité ainsi définie est, comme la mesure, forcément inférieure à l'unité. Nous admettrons l'extension de ces notions de mesure et de probabilité au cas où l'on considère une infinité de segments n'empiétant pas les uns sur les autres, au lieu d'en considérer un nombre limité. Nous remarquerons en outre, que si un ensemble de points est tel que tous ses points sont intérieurs à MN, la probabilité pour qu'un point lui appartienne (c'est-à-dire sa mesure) est au plus égale à la probabilité pour qu'un point appartienne à MN (c'est-à-dire à la mesure de MN).

Ces définitions et remarques étant bien comprises, il va être aisé de démontrer que la mesure de l'ensemble formé des nombres rationnels est égale à zéro, bien que cet ensemble soit dense dans tout l'intervalle OA. Pour le voir, énumérons ces nombres rationnels dans un ordre déterminé, en ayant soin de n'en oublier aucun ; pour fixer les idées, on prendra d'abord les extrémités o et 1 qui sont entiers, puis $\frac{1}{2}$ dont le dénominateur est 2, puis $\frac{1}{3}$ et $\frac{2}{3}$ dont le dénominateur est 3, puis $\frac{1}{4}$ et $\frac{3}{4}$, ensuite $\frac{1}{5}$, $\frac{2}{5}$, $\frac{3}{5}$, $\frac{4}{5}$, et ainsi de suite. Nous allons à chacun de ces points attacher un intervalle ayant ce point pour milieu[1]

1. On négligera les portions de ces intervalles qui seraient extérieures au segment OA ; on négligera aussi les portions de chaque intervalle (ou l'intervalle entier) qui se trouveraient intérieures à un intervalle préalablement construit. Ces détails, sur lesquels je n'insiste pas, ne sauraient infirmer les conclusions, qui restent vraies *a fortiori*.

et par suite renfermant ce point. Les longueurs de ces intervalles seront prises successivement égales à 0,1 ; 0,01 ; 0,001 ; 0,0001 ; etc., leur somme étant ainsi égale à la fraction périodique illimitée 0,11111...., c'est-à-dire à $\frac{1}{9}$. Nous voyons ainsi que l'on peut enfermer tous les nombres rationnels situés sur OA dans des intervalles dont la somme est inférieure à $\frac{1}{9}$, alors que la longueur de OA est l'unité, bien que les nombres rationnels soient denses sur tout OA. Ce résultat est assez paradoxal en apparence pour que j'y insiste un peu. Si l'on se représente matériellement OA comme une tige ou comme un fil et si l'on imagine que la matière soit parfaitement continue, on pourrait concevoir que l'on découpe et enlève les intervalles partiels que nous avons définis. On enlèvera d'abord un intervalle de longueur 0,1 comprenant le point 0, puis un intervalle de longueur 0,01 comprenant le point 1, puis un intervalle de longueur 0,001 comprenant le point $\frac{1}{2}$ et ainsi de suite. Les opérations successives ainsi faites entraîneront la suppression de tous les points rationnels, de sorte que la droite OA sera complètement découpée, en ce sens qu'il n'en subsistera aucune portion continue. Et ce résultat aura été atteint par l'enlèvement de moins de la neuvième partie de la masse de OA; les *portions* non enlevées (si l'on peut les nommer ainsi, bien qu'il ne subsiste aucun *morceau* entier) sont bien plus considérables que les portions enlevées. Remarquons en passant combien notre intuition spatiale du continu est défectueuse, car elle ne nous permet guère de concevoir ce

résultat cependant très simple d'un point de vue arithmétique.

Mais il est clair que rien n'empêcherait de prendre tous nos intervalles 10 fois plus petits, ou 100 fois plus petits, c'est-à-dire de commencer pour le premier par 0,01 ou 0,001 au lieu de commencer par 0,1, chacun restant ensuite 10 fois plus petit que le précédent. La fraction $\frac{1}{9}$ serait alors remplacée par $\frac{1}{90}$ ou $\frac{1}{900}$; il est évidemment loisible d'écrire autant de zéros qu'on le désire à la droite du dénominateur de la fraction ; on arrive donc bien ainsi à la conclusion que l'ensemble des points rationnels a une mesure aussi petite que l'on veut, c'est-à-dire nulle : la probabilité pour qu'un point choisi au hasard sur OA soit rationnel est égale à zéro. La notion géométrique de la mesure conduit au même résultat que la comparaison arithmétique entre les fractions périodiques et les fractions non périodiques.

78. — Lorsque l'on considère, non plus les nombres abstraits de l'arithmétique, mais les nombres concrets définis par l'expérience, la question de savoir s'ils sont rationnels ou irrationnels est au premier abord dépourvue de sens. Un nombre fourni par l'expérience n'est en effet jamais connu exactement, mais seulement à une certaine approximation ; on possède un certain nombre de décimales exactes et on ne sait rien sur les autres décimales ; le nombre peut donc être regardé à volonté comme rationnel ou comme irrationnel. Néanmoins, on a une tendance naturelle, si le résultat d'une expérience prouve qu'un

rapport est compris entre 0,49999 et 0,50001, à admettre qu'il est exactement égal à 0,5 plutôt qu'à une valeur arithmétiquement *plus compliquée*. Pour préciser cette notion un peu vague, il faudrait d'assez longs développements[1] ; on conçoit bien cependant, sans qu'il soit nécessaire d'insister, qu'il est des cas où ce criterium de simplicité amène à des conclusions qui renferment au moins une part de vérité.

79. — L'une des questions les plus importantes de philosophie naturelle dans laquelle intervient la notion de nombre rationnel est l'étude des poids atomiques. On a cru pendant longtemps que les poids atomiques devaient être tous des nombres entiers (ou la moitié de nombres entiers), c'est-à-dire des multiples exacts du poids atomique de l'hydrogène, pris pour unité. Les petites différences étaient attribuées à des erreurs expérimentales. Les progrès réalisés dans la précision des analyses chimiques et dans la préparation des produits purs n'ont pas permis de conserver cette interprétation, qui semble actuellement abandonnée. Que l'on rapporte les poids atomiques à celui de l'hydrogène pris égal à 1 ou à celui de l'oxygène pris égal à 16, les différences entre les poids atomiques et les nombres entiers les plus voisins paraissent supérieures aux erreurs expérimentales. On a néanmoins l'impression, en examinant avec soin une table de poids atomiques, que ces nombres et leurs rapports sont plus voisins de nombres commensurables simples que s'ils étaient choisis au hasard. Pour préciser

1. Voir Emile Borel, Le continu mathématique et le continu physique. *Scientia*, t. VI (1909).

cette impression et rechercher dans quelle mesure on peut présumer que ces coïncidences numériques ne sont pas fortuites, des calculs assez longs seraient nécessaires. Si l'on était conduit, comme il paraît probable, à la conclusion que les rapports de poids atomiques peuvent être rattachés à des nombres rationnels dont ils diffèrent très peu, il y aurait lieu de se demander si l'on ne pourrait pas expliquer d'une part, la valeur rationnelle exacte, et d'autre part, la petite différence.

80. — La théorie de la mesure dont nous avons indiqué le principe pour les ensembles linéaires, s'étend sans difficulté aux ensembles situés dans le plan et dans l'espace à trois ou à un plus grand nombre de dimensions. Sous cette forme générale elle peut être très utile pour l'application de considérations statistiques aux questions de stabilité qui se posent pour les problèmes mécaniques les plus généraux.

Dans le domaine des mathématiques pures, c'est la mesure des ensembles plans qui a été le plus utilisée, car elle trouve son application immédiate dans la théorie des fonctions d'une variable complexe, théorie dont la place est si grande en analyse, depuis les travaux de Cauchy.

On sait que l'étude des problèmes analytiques les plus importants se trouve à la fois simplifiée et éclaircie lorsqu'on envisage ces problèmes, non pas seulement pour les valeurs réelles de la variable, mais aussi pour les valeurs complexes. Dans le langage géométrique, une variable réelle est représentée par un point d'une droite, tandis qu'une variable complexe est représentée par un point d'un plan.

L'étude des fonctions d'une variable complexe est donc l'étude de certaines fonctions qui dépendent d'une manière particulière[1] de la position d'un point variable dans un plan indéfini : ce plan est dit le champ complexe, par opposition à la droite indéfinie, qui est le champ réel.

Lorsque l'on étudie les propriétés d'une fonction dans le champ complexe, on constate que certains points que l'on appelle points singuliers et parfois certaines régions que l'on appelle régions singulières, jouent un rôle très important. Il arrive souvent qu'il est possible d'établir pour la fonction étudiée certaines propriétés simples, à condition d'exclure certaines régions voisines des points singuliers ou des régions singulières. Dans le cas où les régions ainsi exclues ont une mesure qui est une fraction connue de la région totale dans laquelle on considère la fonction, il est clair que l'on connaît la probabilité pour que, un point étant choisi au hasard dans cette région totale, ce point ne se trouve pas dans les régions exclues et que par suite la fonction y possède les propriétés simples que l'on a établies. On peut dire, pour abréger, que l'on connaît la probabilité pour que la fonction possède ces propriétés.

Dans certains cas, soit que la région totale considérée soit infinie, soit que les régions exclues puissent être rendues de mesure aussi petite que l'on veut, on sera conduit à dire que la probabilité est égale à un, ou du moins

1. Particulière, à cause de la condition fondamentale de monogénéité introduite par CAUCHY : existence d'une dérivée unique.

infiniment voisine de un, tandis que la probabilité contraire est infiniment petite. Ce langage est souvent très commode et son emploi systématique peut conduire parfois à des résultats importants. Il n'est guère possible de donner ici des exemples de cette *méthode d'exclusion ;* mais j'ai tenu à signaler tout au moins cet emploi du langage et des méthodes de la théorie des probabilités dans des questions d'analyse pure.



TROISIÈME PARTIE

LA VALEUR DES LOIS DU HASARD

CHAPITRE VIII

LA VALEUR PRATIQUE DES LOIS DU HASARD

81. Retour sur les principes et les définitions. — 82. Le jeu équitable. — 83. L'espérance relative et l'espérance mathématique dans les loteries. — 84. L'utilité du calcul dans les problèmes de jeux de cartes. — 85. L'utilité du calcul est la même dans toute question pratique. — 86. L'abus et le mépris des chiffres. — 87. La distinction entre la probabilité objective et la probabilité subjective. — 88. La vie pratique et le hasard. — 89. L'unité humaine de probabilité très petite. — 90. Le « scepticisme » de Poincaré. — 91. La sensibilité individualiste et les plaisanteries sur la statistique. — 92. Les lois du hasard et la vaccination. — 93. La portée sociale des lois du hasard. — 94. Les mathématiques sociales et les excès de l'individualisme. — 95. La morale et les coefficients d'altruisme. — 96. Les coefficients d'altruisme peuvent-ils être négatifs ? — 97. Le patriotisme et l'internationalisme.

81. — Nous avons essayé, dans la première partie de cet ouvrage, d'exposer comment on est arrivé à soumettre au calcul les lois du hasard ; nous avons ensuite, dans la seconde partie, passé rapidement en revue les applications pratiques et scientifiques des méthodes ainsi créées ; nous devons maintenant nous demander quelle est la valeur pratique, scientifique et philosophique de ces applications ; doit-on considérer cette valeur comme diminuée par ce je ne sais quoi de mystérieux qu'évoque le mot de « hasard », ou doit-on au contraire penser que c'est l'étude approfondie des lois du hasard qui nous rensei-

gnera le mieux sur la valeur de toute connaissance humaine ?

Il me paraît nécessaire, pour aborder ces questions essentielles, de reprendre le problème du hasard en quelque sorte *ab ovo*, en regardant les chapitres précédents comme une simple introduction destinée à étayer d'exemples précis les réflexions qui vont suivre. C'est dans cet esprit que nous allons, dans ce chapitre, étudier tout d'abord la valeur de la probabilité dans la vie pratique et rechercher ensuite comment la sensibilité individualiste fait souvent obstacle à l'acceptation par beaucoup d'hommes de conclusions qui s'imposent cependant à leur raison ; nous terminerons par de brèves indications sur le rôle que pourrait jouer la théorie du hasard dans une morale basée sur la solidarité et dans l'évaluation de la valeur sociale des individus.

Le calcul des probabilités est l'étude des lois du hasard.

On a déjà fait remarquer que cette définition n'explique une contradiction que par une autre contradiction. Si l'on ne comprend pas comment l'on peut parler de *calcul* à propos de *probabilité*, on comprendra encore moins qu'il puisse être question de *lois* à propos du *hasard*. Le hasard, n'est-ce pas précisément ce qui est en dehors de toute règle, de toute loi ? Et l'expérience de chaque jour ne nous apprend-elle pas à nous défier des lois auxquelles on prétend soumettre les événements fortuits ? Ne se produit-il pas fréquemment des coïncidences bizarres, des accidents étranges, contraires en apparence à toute probabilité ?

« De tels événements se produisent sans

doute, affirmera le mathématicien, mais moins fréquemment qu'il ne le semble, et leur fréquence même est réglée par les lois que leur arrivée semble démentir. On peut d'ailleurs donner de la justesse du calcul des probabilités d'excellentes preuves financières : une compagnie d'assurances bien gérée fait toujours des bénéfices et la roulette n'a jamais ruiné son gérant. »

« Mais alors, répondra le contradicteur, à propos de ce dernier exemple, vous regardez donc comme impossible que la roulette soit ruinée? Il suffirait cependant que je sois seul joueur et que je joue constamment le numéro qui va sortir. Il est sans doute bien peu probable que j'aie un tel flair, mais cela est évidemment possible à concevoir. Que deviennent alors vos principes? »

Ce débat pourrait se poursuivre longtemps, car la difficulté qui en fait le fond ne peut être élucidée que si on l'aborde nettement, de face, en commençant par étudier la notion même de probabilité, avant de légitimer l'application que l'on y fera du calcul.

82. — Considérons d'abord la probabilité dans le jeu le plus simple, celui de *pile ou face*. On jette en l'air une pièce de monnaie et le pari porte sur le côté qui sera apparent après la chute.

Pour éviter toute difficulté provenant de la dissymétrie de la pièce, nous pouvons supposer que l'on a tout simplement un jeton métallique, dont les deux faces sont absolument pareilles et ne se distinguent l'une de l'autre que par une différence de teinte. Il n'y a alors aucune raison pour que le côté visible après la chute soit l'un

plutôt que l'autre, surtout si l'on suppose que celui qui lance la pièce a les yeux fermés et ne peut, par suite, distinguer les deux côtés. Il y a donc chances *égales* pour le côté que l'on appellera *pile* et pour celui que l'on appellera *face*. On convient de dire que la probabilité de chacun d'eux est *un demi*.

Si l'on est obligé de parier ou décidé à le faire, il n'y a pas de raison de parier pour pile plutôt que pour face; si les mises sont égales, le jeu est *équitable;* si les mises sont inégales, il est *avantageux* pour celui des joueurs dont la mise est la plus faible.

De ces constatations, évidentes par le simple bon sens, pouvons-nous tirer une règle de conduite?

Avant de répondre à cette question, remarquons tout d'abord que nous laissons ici de côté les *raisons morales* contre le jeu; elles sont en dehors de notre cadre, et nous nous plaçons systématiquement au point de vue utilitaire. Mais, même en simplifiant ainsi le problème, ce n'est que dans des cas très particuliers que l'étude faite sur la probabilité au jeu de pile ou face *suffit* à nous donner une règle de conduite. On peut avoir d'excellentes raisons pour ne pas jouer un jeu équitable ou même avantageux; on peut, quoique plus rarement, avoir de bonnes raisons pour jouer un jeu désavantageux.

Pierre possède exactement un million; on lui propose de le jouer à pile ou face, contre un million (jeu équitable), ou même contre un million et cinquante francs (jeu théoriquement avantageux). Il est clair qu'à moins de circonstances très particulières, il est de l'intérêt de Pierre de refuser.

Jacques se trouve isolé, sans relations, dans une contrée éloignée ; il est fort riche et recevra demain une somme importante ; mais il a très grand intérêt à prendre un paquebot qui part dans une heure et il n'a sur lui que 300 francs, alors que la traversée à payer d'avance coûte 400 francs. On lui propose de jouer à pile ou face ses 300 francs contre 200 (jeu théoriquement désavantageux) ; il a évidemment intérêt à accepter.

Si, au lieu d'un jeu simple comme pile ou face, on considère un jeu un peu plus compliqué, comme la loterie, il arrive fréquemment que des considérations étrangères au simple calcul de la probabilité interviennent pour faire adopter telle ou telle règle de conduite. Par exemple, on ne peut pas affirmer qu'il soit absolument déraisonnable de payer 2 francs un billet d'une loterie, où il y a un million de billets et un lot d'un million. La personne unique qui gagne le million trouve même généralement qu'en prenant son billet elle a été sage et prévoyante.

Cependant, elle a payé 2 francs un billet qui, mathématiquement, n'en valait qu'un.

Inversement, supposons que Jacques possédant exactement 100.000 francs, on lui offre 100 francs à condition que si un numéro déterminé sort le premier à une loterie de 2.000 billets, il fera abandon de toute sa fortune. La proposition est mathématiquement avantageuse ; elle est pratiquement inacceptable.

83. — On a proposé une méthode empirique pour tenir compte du fait que l'importance d'une même somme d'argent dépend de la fortune de

son propriétaire; la méthode consiste à admettre que c'est le rapport de la perte ou du gain probable à la fortune totale du joueur qui constitue son *espérance relative*. Nous ne nous attarderons pas à discuter cette convention, dont Joseph Bertrand a montré tout l'arbitraire ; remarquons simplement que le chiffre de la fortune que possède un homme est loin d'être le seul motif qui le rende sensible à un gain ou à une perte déterminée.

Le seul moyen de ne pas être conduit à des considérations aussi conventionnelles, est de renoncer à parler de ce qu'on appelle l'*espérance mathématique*, ou tout au moins à ne regarder cette expression que comme désignant une quantité qu'il est souvent avantageux d'introduire dans les calculs de probabilités. Mais il ne faut pas la prendre au pied de la lettre ; on fait cadeau à Pierre d'un billet d'une loterie dont le lot unique est de 100.000 francs pour un million de billets ; son espérance mathématique est le quotient de la valeur du lot par le nombre de billets, c'est-à-dire 10 centimes. Et cependant, s'il est pessimiste et mourant de faim, il préférera de beaucoup qu'on lui donne deux sous pour acheter du pain, tandis que s'il a de l'imagination, il sera pour un jour presque aussi content que si on lui avait donné les 100.000 francs. En évaluant son espérance mathématique, on a fait un calcul, sans doute exact, mais sans valeur pratique.

Les organisateurs de loterie le savent bien ; ils cherchent à frapper l'imagination du public. On vendra plus facilement 20 francs un billet auquel est attaché une espérance mathématique de 8 francs, si la répartition des lots est savamment calculée, qu'un billet auquel est attachée une espérance

mathématique de 12 francs, mais qui ne peut faire gagner aucun gros lot sensationnel.

Concluons que la probabilité et la certitude n'ont point de commune mesure : avoir une chance sur deux de posséder 1.000 francs vaut mathématiquement 500 francs ; tel préférera avoir sûrement 400 francs en poche, pensant qu'il vaut mieux tenir que courir et tel autre, au contraire, payera volontiers cette chance 600 francs.

Allons-nous donc être conduit à dénier toute valeur pratique à la théorie des probabilités, en raison de cette impossibilité d'évaluer la probabilité au moyen de la certitude ?

84. — S'il n'est pas possible de trouver une commune mesure entre la probabilité et la certitude, il nous reste la ressource de chercher à évaluer la probabilité au moyen de la probabilité elle-même, c'est-à-dire au moyen d'une quantité *de même nature*. La théorie des probabilités sera ainsi l'*ensemble des procédés mathématiques qui permettent de définir d'une manière simple une probabilité équivalente à une probabilité donnée d'une manière compliquée*. Éclaircissons cette définition par quelques exemples.

En voici un où le procédé mathématique se réduit à un calcul très simple. Pierre et Paul jouent une série de parties de pile ou face ; on note le résultat de chaque partie ; pendant que se jouent les neuf premières, aucun des joueurs ne doit rien payer à l'autre ; à partir de la dixième, Paul verse 1 franc à Pierre avant chaque partie et, en échange, Paul s'engage à verser une certaine somme à Paul après chaque série de dix

parties qui auront toutes les dix amené face. *Quelle doit être cette somme pour que la situation de Paul soit la même que s'il prenait des billets de 1 franc à une loterie équitable?* Un calcul aisé dont on trouvera les éléments dans tous les traités de calcul de probabilités donne comme réponse la dixième puissance de deux, c'est-à-dire 1.024 francs. Si Pierre s'engage à verser 500 francs à Paul en cas de réussite, et si les joueurs ne font que dix parties, c'est-à-dire si Paul ne risque que 1 franc, sa situation est la même que s'il prenait un billet de 1 franc à une loterie de 1.024 billets dont le lot unique serait de 500 francs. Doit-il le faire ou non ? Le calcul ne peut répondre à cette question, mais seulement éclairer Paul, en précisant la nature de la chance qu'il court ; c'est à lui de décider ensuite s'il juge opportun de la courir.

Le calcul précédent était de ceux que toute personne un peu habituée aux problèmes de probabilité fait d'une manière presque réflexe, sans s'apercevoir qu'elle résout un problème. Il est des cas où le calcul nécessaire exige forcément un temps assez long, même avec l'emploi des ressources de l'arithmétique et de l'analyse. C'est ce qui se passe dans la plupart des problèmes qui se posent dans les jeux de cartes et c'est là sans doute une des raisons de l'attrait de ces jeux pour ceux qui s'y adonnent : il faut, à chaque instant, résoudre des problèmes compliqués dont la solution rigoureuse serait fort longue ; le joueur est obligé de trouver rapidement une solution par des procédés empiriques, et n'est pas toujours certain que cette solution est la meilleure.

Pour donner une idée de la complication des calculs auxquels conduisent les problèmes de jeux de cartes, nous allons choisir un des exemples les plus simples, dans un jeu qui est lui-même un des plus simples[1].

Pierre et Paul jouent à l'écarté. Paul a donné et a retourné le huit de carreau; Pierre a le sept de carreau, le roi, la dame, le valet et l'as de pique. Quelles sont les chances de gain de Pierre en jouant d'autorité ?

Nous précisons le problème en ajoutant l'hypothèse que Paul joue pour plus d'un point (sinon on devrait admettre qu'il n'a pas le roi d'atout); les cartes de Paul sont alors 5 cartes choisies parmi les 26 cartes, autres que celles du jeu de Pierre et la retourne, à savoir : 6 atouts, 4 piques, 8 carreaux et 8 trèfles.

Quelle est la probabilité pour que Paul n'ait pas d'atout? Pour la calculer, on doit calculer d'abord le nombre total des jeux possibles, c'est le nombre des combinaisons de 26 objets 5 à 5 ; et ensuite le nombre des jeux sans atouts, c'est le nombre des combinaisons de 20 objets 5 à 5. En divisant ce dernier nombre par le premier, on trouve, pour la probabilité cherchée, 0,2357 environ, c'est-à-dire qu'il y a 2.357 chances sur 10 000 pour que Paul n'ait pas d'atout; on trouve de même qu'il y a 4.419 chances sur 10.000 pour qu'il en ait un seul, 3.224 sur 10.000 pour qu'il en ait 2, 3, 4 ou 5.

Il est clair que si Paul n'a pas d'atout, Pierre fait la vole et que, s'il a un seul atout, Pierre gagne le point.

Si Paul a plus d'un atout, Pierre perd, à moins que Paul n'ait précisément 2 atouts et 3 piques : on trouve que la probabilité de ce cas est 9 sur 10.000.

En résumé, il y a 2.357 chances sur 10.000 pour que Pierre gagne deux points; 4.428 (4.416 + 9) chances sur 10.000 pour qu'il gagne un point, et 3.215 chances sur 10.000 pour qu'il perde deux points (ayant joué d'autorité). D'une manière approximative mais plus simple, on peut dire qu'il y

1. Je laisse de côté le baccara, jeu rudimentaire, dont la théorie mathématique peut être faite complètement et qui, pour cette raison, est un jeu sans intérêt propre ; c'est seulement un moyen rapide de gagner ou de perdre de l'argent.

a, à peu près, 2 chances sur 9 pour que Pierre gagne deux points, 4 chances sur 9 pour qu'il en gagne un et 3 chances sur 9 pour qu'il en perde deux.

Voilà les renseignements que le calcul donne à Pierre; c'est ensuite à lui de se décider. Il semble que, s'il joue une partie ordinaire, sa décision ne sera pas douteuse : en effet, sur 9 parties analogues, il peut espérer gagner en tout 8 points et en perdre 6[1]; mais supposons qu'un spectateur qui suit le jeu de Pierre lui propose un pari très important sur le gain du coup; il sera facile à Pierre, à l'aide des chiffres précédents, de reconnaître si les conditions du pari sont équitables, avantageuses ou désavantageuses et il aura ainsi *l'un* des éléments de sa décision quant à l'acceptation de ce pari.

85. — La valeur pratique de la théorie des probabilités nous apparaît donc comme relative; le problème pratique à résoudre est simplifié dans des termes, mais n'est pas modifié dans son essence; il reste un problème de probabilité.

Il n'est peut-être pas inutile de faire observer que ce caractère relatif n'est pas spécial à l'application des mathématiques aux probabilités; on le retrouve dans toutes les applications pratiques des mathématiques, bien qu'on soit parfois tenté de leur attribuer une valeur absolue.

1. Nous avons laissé de côté un point important; nous avons raisonné comme si, dans le cas où les deux joueurs demandent des cartes, les chances de gain étaient égales pour l'un et l'autre; ici, on verrait qu'elles sont plus faibles pour Pierre que pour son adversaire, ce qui est une raison de plus pour que Pierre joue d'autorité; mais, d'une manière générale, pour traiter un problème d'écarté on doit se demander si le fait de demander des cartes augmente ou diminue les chances de gain (en tenant compte aussi de la probabilité pour que l'adversaire en refuse).

Je désire éclairer une salle ; on me donne les éléments nécessaires pour calculer le prix de revient de l'éclairage au gaz et à l'électricité. Après des calculs plus ou moins longs, je trouve que je dépenserai 30 francs par mois avec le gaz, et 32 francs avec l'électricité. Ce résultat précis sera l'un des éléments de ma décision ; je suis mieux renseigné que si je connaissais seulement le prix du mètre cube de gaz et celui de l'hectowatt-heure ; mais le calcul ne m'impose pas ma décision.

Il n'est pas nécessaire de multiplier les exemples pour se rendre compte du rôle du calcul dans la vie pratique ; il met sous une forme plus facile à saisir certains éléments de nos décisions ; il remplace certaines données, plus ou moins complexes, par un petit nombre de chiffres simples.

Il est cependant des cas où le calcul paraît suffire à déterminer notre décision ; mais cela tient à ce que nous nous étions fixés d'avance une certaine règle de conduite ; par exemple, dans l'exemple de l'éclairage au gaz ou à l'électricité, si j'ai décidé, avant tout calcul, de choisir l'éclairage le meilleur marché, le calcul m'impose le choix du gaz. Si j'avais décidé de choisir l'électricité, pourvu que la dépense supplémentaire ne dépassât pas 5 francs par mois, le calcul m'imposerait le choix de l'électricité.

De même, je puis avoir l'intention de consacrer une certaine somme à l'achat de billets de loterie, et être décidé à choisir, entre deux loteries qu'on me propose, ou bien celle qui me donnera l'espérance mathématique la plus forte, ou bien celle qui me donnera la plus grande chance de gagner un lot supérieur ou égal à 100.000 francs.

Dans chaque cas, la théorie des probabilités me dictera ma décision.

86. — Malgré l'évidence de ces remarques, j'ai cru devoir y insister un peu, car l'intervention du calcul dans les décisions de la vie pratique donne trop souvent lieu à l'un de deux jugements extrêmes ; pour les uns, il est absurde de mêler le calcul à une décision dont certains éléments ne sont pas exprimables en chiffres ; pour d'autres, les chiffres ont une vertu magique qui rend infaillibles tous ceux qui les emploient suivant les règles.

Ces deux tendances opposées correspondent d'ailleurs, au fond, à un même état d'esprit ; c'est parce que les chiffres leur paraissent avoir une valeur absolue éliminant toute discussion que certains esprits redoutent leur intervention et préfèrent se passer de leur secours que de s'assujettir à leur joug.

On dit volontiers : rien n'est exact comme un chiffre, et l'on dit aussi : rien n'est brutal comme un chiffre. Les uns sont décidés à se soumettre aveuglément au chiffre exact ; les autres refusent d'être mis en présence du chiffre brutal.

L'origine de cette illusion est aisée à indiquer : on ne peut faire un calcul que sur des chiffres précis ; il est tout au moins fort long de rechercher les diverses solutions d'un problème qui correspondent à des valeurs diverses de données imprécises, et l'on recule généralement devant cette tâche. On est alors conduit à adopter une valeur déterminée pour chacun des éléments du calcul, même si cet élément est mal connu. C'est ainsi qu'on établit un devis ; on dit à chaque ins-

tant : il faudra pour telle tâche de quatre à six journées d'ouvrier ; *mettons* cinq journées. Il n'est pas jusqu'à l'imprévu qu'on ne calcule rigoureusement ; d'après des données empiriques, on évalue à forfait les frais imprévus une certaine fraction, le dixième par exemple, du devis primitif.

Sur ces données précises, on peut faire des calculs précis, qui conduisent à un résultat précis. Et, plus les calculs sont longs, plus le résultat risque d'être inexact et plus cependant on a le temps d'oublier que les données étaient imprécises et de se laisser aller à la confiance qu'inspire l'exactitude des opérations arithmétiques correctement faites.

La même illusion se produit fréquemment dans les statistiques ; on lit, chaque année, dans les journaux, que la production totale du blé en France est évaluée, par exemple, à 115.200.000 quintaux. Ce chiffre a été obtenu en additionnant un grand nombre d'indications particulières, dont chacune était inexacte ; sa masse inspire cependant une certaine confiance, et on en tire volontiers des déductions économiques.

Gardons-nous donc d'avoir trop de confiance dans les chiffres ; ce sera le meilleur moyen d'éviter l'excès contraire, qui consiste à repousser complètement leur aide. Il faut s'en servir, mais en n'oubliant jamais qu'ils ne créent point la certitude : le calcul des probabilités apparaît alors comme aussi justifié que tout autre calcul ; sa valeur pratique est exactement la même.

87. — Il est cependant certaines applications pratiques de la théorie des probabilités, dont les

résultats palpables sont tellement visibles pour tous que l'on est tenté de se demander si nous n'avons pas été trop sévères en affirmant qu'un calcul portant sur des probabilités ne peut pas conduire à la certitude.

Les actions d'une compagnie d'assurances sérieuses, celles même du cercle de Monaco, sont un placement dont le revenu est très régulier. Il n'y a pas de rapport entre la conduite d'un capitaliste qui achète ces actions et celle d'un joueur qui risque tout son argent au baccara ou à la loterie.

Cette différence est assez grande pour que l'on ait pu être amené à parler de la valeur *objective* du calcul des probabilités :

> Un joueur veut tenter un coup; il me demande conseil. Si je le lui donne, je m'inspirerai du calcul des probabilités mais je ne lui garantirai pas le succès. C'est là ce que j'appellerai la *probabilité subjective*... Mais je suppose qu'un joueur assiste au jeu, qu'il en note tous les coups et que le jeu se prolonge longtemps; quand il fera le relevé de son carnet, il constatera que les événements se sont répartis conformément aux lois du calcul des probabilités. C'est là ce que j'appellerai la *probabilité objective*... Il existe de nombreuses sociétés d'assurances qui appliquent les règles du calcul des probabilités et elles distribuent à leurs actionnaires des dividendes dont la réalité objective ne saurait être contestée. (Poincaré, *Science et hypothèse*.)

Si j'ai tenu à citer ce passage, c'est qu'il me paraît nécessaire de prévenir une confusion, que n'a certainement pas faite Poincaré, mais que pourrait faire un lecteur non averti. Ce n'est pas une différence de nature qui sépare la probabilité objective de la probabilité subjective, mais seulement une différence de degré. Un résultat du calcul des probabilités mérite d'être appelé objectif,

lorsque sa probabilité devient assez grande pour se confondre pratiquement avec la certitude. Il importe peu alors qu'il s'agisse de prévoir des phénomènes futurs ou de recenser des phénomènes passés ; on peut également affirmer que la loi sera ou a été vérifiée.

Il y a, par exemple, à Paris, environ 1.000 naissances par semaine, et l'on peut admettre comme un fait d'expérience que la probabilité de la naissance d'un garçon est égale à la probabilité de la naissance d'une fille [1].

Est-il possible que pendant une semaine déterminée (passée ou future) il ne naisse à Paris que des garçons? La probabilité d'un tel événement est à peu près égale à la probabilité d'amener mille fois de suite pile au jeu de pile ou face ; il y a une chance favorable contre un nombre de chances défavorables égal à la millième puissance de deux, c'est-à-dire à un nombre d'environ *trois cents chiffres*. Cette chance favorable est à peu près celle que l'on aurait d'obtenir les deux premiers vers du *Cid* en tirant au sort dans un chapeau les lettres de l'alphabet français : on y aurait mis les vingt-cinq lettres, on en tirerait une, que l'on inscrirait et que l'on remettrait dans le chapeau ; on en tirerait une seconde après

1. En réalité, l'égalité n'est pas exacte, et il n'est pas non plus absolument exact d'assimiler, au point de vue de la loi des écarts, les naissances à un jeu de pile ou face. Mais les approximations que nous faisons sont parfaitement légitimes. notre but étant de donner un exemple et non de faire un calcul rigoureux. On trouvera une intéressante discussion de la probabilité des naissances masculines et féminines dans le *Traité du calcul des probabilités*, de Joseph Bertrand (Gauthier-Villars). Voir aussi Emile Borel, *Eléments de la théorie des probabilités* (Hermann) et E. Carvallo : *Le calcul des probabilités et ses applications* (Gauthier-Villars).

avoir agité et ainsi de suite. Il n'est évidemment pas *impossible*, au sens rigoureux du mot, que l'on obtienne précisément les deux premiers vers du *Cid*, en procédant ainsi. Cela nous paraît cependant tellement peu vraisemblable que, si l'expérience faite devant nous réussissait, nous serions absolument persuadés qu'elle était truquée.

Nous n'avons pas à revenir ici, sur les principes au moyen desquels on arrive à calculer ces probabilités complexes qui approchent de la certitude, en partant de probabilités simples beaucoup plus faibles. Sachant que la probabilité d'amener *face* est un sur deux, on trouve que la probabilité pour que l'on amène moins de 400.000 fois *face* sur un million de coups est absolument négligeable. Dans cet exemple, le calcul est simple et les bases sur lesquelles il repose ne paraissent pas contestables ; les choses sont plus compliquées dans les problèmes d'assurances, où interviennent des probabilités empiriques sur la vie humaine. La vérification objective des résultats du calcul peut alors être considérée comme apportant aux principes employés une confirmation nécessaire : c'est en ce sens qu'il faut entendre les remarques de Poincaré, que nous avons partiellement citées tout à l'heure.

88. — Ce n'est pas seulement dans les problèmes de probabilité, c'est dans tous les actes de la vie que nous agissons comme si une probabilité très grande équivalait à une certitude.

Lorsque nous disons que les bénéfices d'une compagnie d'assurances dépendent de l'exactitude du calcul des probabilités, nous pensons aux

calculs qu'ont faits les actuaires de cette compagnie pour l'établissement de ses tarifs. Mais il ne suffit pas que les tarifs soient bien établis pour que la Compagnie fasse des bénéfices, il faut surtout qu'elle ait une clientèle, et la régularité de la clientèle est une question de probabilité qui intervient dans toutes les entreprises.

Mais nul n'est tenu d'être actionnaire ; n'est-il pas possible de limiter notre activité de manière que la probabilité n'y joue aucun rôle ? Il est aisé de voir qu'une telle prétention mènerait tout droit à la folie.

Il suffit pour cela d'imaginer la vie d'un homme dans ces conditions : il rejettera de son alimentation toutes les matières suspectes de pouvoir renfermer des poisons ou des microbes pathogènes, ce qui sera déjà bien compliqué ; il vivra dans une maison incombustible, ne sortira pas dans les rues quand il fait du vent de peur qu'une tuile ne l'écrase, etc.

Mais ce n'est pas tout : il reçoit une dépêche signée du nom d'un de ses amis et se rapportant à une question importante qu'eux seuls connaissent : qui lui assure que cette dépêche vient bien de son ami ? N'y a-t-il pas une erreur de transmission, une coïncidence qui a transformé complètement l'adresse, le texte et la signature ; la probabilité pour qu'il en soit ainsi pourrait être calculée et ne serait pas rigoureusement nulle.

La probabilité n'est pas nulle non plus pour que l'on confonde une personne inconnue avec un ami, que cette personne inconnue vous confonde en même temps avec un de ses amis, vous interpelle par votre nom (qui serait comme par hasard le même), vous parle de plusieurs amis

communs (ayant aussi les mêmes noms ou prénoms) et finisse par vous emprunter de l'argent sous un prétexte qui paraît très naturel.

Il est inutile d'insister ; concluons que chacune de nos décisions est basée sur des données simplement probables, mais assez probables pour que nous agissions comme si elles étaient certaines. Plusieurs des données que nous fournit la théorie des probabilités proprement dite sont assez probables pour pouvoir être rangées dans cette catégorie. La valeur pratique de la théorie des probabilités est, dans ce cas, exactement la même que celle de toutes nos autres connaissances.

89. — Nous pourrions en rester là, mais il est une question importante dont il paraît nécessaire de dire un mot, bien qu'elle ne soit pas d'ordre mathématique, mais d'ordre purement psychologique et ne soit par suite susceptible d'aucune solution vraiment précise. Mais elle se rattache trop intimement à ce qui précède et vient trop naturellement à l'esprit pour que nous puissions refuser de l'aborder.

On peut la formuler ainsi : *à partir de quelle limite la probabilité peut-elle être confondue avec la certitude?* Il est évident que l'on ne peut faire à une question ainsi posée une réponse rigoureuse et exacte. Il serait absurde de dire qu'il y a certitude lorsqu'il y a une chance défavorable sur 1.000.000, tandis qu'il y a probabilité lorsqu'il y a une chance défavorable sur 999.999. Le vieux sophisme du *tas de blé* trouve ici son application.

Mais d'autre part, cela a-t-il un sens de dire

qu'on confond avec la certitude une *très grande* probabilité[1] ?

Il est clair que le mot très grand n'a et ne peut avoir aucun sens *absolu* ; il a seulement un sens *relatif*. Un milliard est-il un nombre très grand ? S'il s'agit de francs à recevoir, on jugera généralement que c'est beaucoup ; s'il s'agit de litres d'eau à verser dans l'océan il est certain que c'est très peu. Les mots *très grand* et *très petit* n'ont donc aucun sens, si l'on ne précise pas l'unité de comparaison.

Où pouvons-nous chercher cette unité de comparaison pour les probabilités ?

Il ne semble pas que l'on puisse faire deux réponses à cette question : nous devons nous rappeler que nous sommes des hommes et chercher une *unité humaine*. On peut d'ailleurs, pour cela, se placer à des points de vue divers, qui conduisent à des résultats sensiblement analogues.

Nous pouvons porter notre attention sur la durée de la vie humaine. Cette durée ne dépasse guère 30.000 jours ; un homme qui accomplit un acte plusieurs fois par jour, pendant toute sa vie, arrive à l'accomplir quelques centaines de milliers de fois ; s'il s'agit d'un acte très simple, comme de faire un pas, d'avaler une bouchée, ou de respirer, ce pourra être au plus quelques centaines de millions de fois.

Si, dans un tel acte, une circonstance quelconque peut éventuellement présenter des inconvénients graves, nous chercherons à les éviter, si nous les connaissons, ce qui arrivera si ces incon-

1. En langage mathématique, il faudrait dire *une probabilité très voisine de un*. Nous parlons ici le langage vulgaire et prenons comme valeur de la probabilité le nombre des chances favorables qu'il y a contre une chance défavorable.

vénients se présentent effectivement assez souvent pour que nous ayons pu les remarquer. Mais, si la probabilité de ces inconvénients est assez faible pour qu'ils soient généralement ignorés, la plupart des hommes n'en tiendront pas compte. On est ainsi amené à admettre comme un fait que, dans la conduite habituelle de la vie, nous négligeons généralement les probabilités inférieures à un millionième (sauf quand nous prenons un billet de loterie).

On arrive à la même conclusion en considérant un autre élément humain, le nombre des hommes qui vivent dans une ville ou un pays. Il y a à Paris moins d'un million d'hommes adultes ; les journaux rapportent chaque jour des accidents ou incidents bizarres arrivés à l'un d'eux ; la vie serait impossible si chacun craignait continuellement pour lui-même toutes les aventures qu'on peut lire dans les faits divers ; cela revient à dire qu'on doit négliger pratiquement les probabilités inférieures à un millionième. Telles sont les probabilités pour qu'il se produise des coïncidences invraisemblables ; lorsqu'elles sont présentées d'une manière systématique, elles excitent le rire : c'est un procédé bien connu des vaudevillistes et des fabricants de films comiques.

Je sais bien qu'un homme sage et prudent dira : il faut toujours mettre de son côté le plus de chances possible ; si je sais qu'en agissant de telle manière, j'ai, ne serait-ce qu'une chance sur un milliard d'être tué, tous les raisonnements ne me décideront pas à courir gratuitement cette chance.

Sans doute, ô homme sage et prudent ; mais toute votre prudence ne vous empêchera pas de

courir journellement des dangers dont la probabilité est bien plus forte :

> Souvent la peur d'un mal fait tomber dans un pire.

Pour savoir distinguer le pire, il est bon de connaître les probabilités des diverses éventualités ; c'est en ce sens que la théorie des probabilités intervient d'une manière plus ou moins consciente dans toutes nos décisions ; il y aurait souvent avantage à rendre son intervention plus consciente en la précisant par des chiffres, à condition toutefois de ne pas oublier que, si nous pouvons prendre les chiffres comme conseillers, nous ne devons jamais être leur esclave.

90. — La théorie des probabilités a été créée et perfectionnée par des hommes qui furent à la fois des savants illustres et des penseurs universellement admirés : qu'il suffise de citer les noms de Pascal, Buffon, d'Alembert, Condorcet, Euler, Laplace, parmi les plus célèbres ; il est singulier de constater que, malgré l'autorité qui s'attache à de tels noms, on rencontre à chaque instant, chez les personnes incapables d'approfondir les théories mathématiques, des doutes ou même des négations sans l'ombre d'une preuve au sujet des résultats les plus solidement établis. C'est là un phénomène des plus curieux, car il contraste singulièrement avec la foi aveugle et parfois excessive que le public accorde généralement aux affirmations des « savants ».

Je ne crois pas qu'il faille en chercher la raison dans le « scepticisme transcendantal » auquel se sont parfois laissés aller quelques-uns des penseurs les plus éminents qui ont écrit sur le calcul

des probabilités, — je fais particulièrement allusion à Joseph Bertrand et surtout à Henri Poincaré. Lorsque celui-ci discute les fondements de la connaissance, il est porté à contester également la *valeur absolue* des opinions classiques sur le mouvement de la terre et sur les lois du hasard; mais c'est là une opinion métaphysique sans plus de valeur pratique que la négation philosophique de l'existence du monde sensible — la comparaison est de Poincaré lui-même. C'est seulement en les comprenant mal qu'on a pu croire que Bertrand ou Poincaré ne regardent pas les résultats de la théorie des probabilités comme des vérités scientifiques aussi certaines que toute vérité scientifique. Mais ce n'est pas parmi leurs lecteurs que se trouvent les sceptiques — ou plutôt les négateurs — dont je voulais parler ; ceux-ci n'ont pas généralement la culture mathématique nécessaire pour lire de tels ouvrages ; ce n'est pas leur raison, c'est leur sensibilité qui est choquée par les conclusions de la théorie des probabilités, ou du moins par la manière dont ils les comprennent et les interprètent.

Nous sommes donc conduits à nous demander ce qui, dans la théorie des probabilités, heurte à tel point la sensibilité de beaucoup d'hommes que ceux-là seulement l'acceptent pour vrai, dont la raison est assez forte pour suivre dans tous leurs détails les déductions logiques; pour les autres, l'argument d'autorité, si puissant lorsque leur sensibilité est indifférente et impartiale, ne suffit pas à vaincre leurs répugnances.

En essayant d'approfondir les raisons pour lesquelles la théorie des probabilités est antipathique à beaucoup d'esprits, j'espère arriver à faire voir

que cette antipathie repose en grande partie sur un malentendu ; il serait désirable que ce malentendu soit dissipé, car la vulgarisation des conclusions, sinon des méthodes, de cette branche de la science, serait d'une grande utilité sociale.

91. — Chacun de nous tient particulièrement à tout ce qui constitue son individualité et participe ainsi plus ou moins à cette *sensibilité individualiste* que l'on a parfois signalée comme l'apanage de quelques esprits d'élite.

Le trait dominant de la sensibilité individualiste est en effet celui-ci ; le sentiment de la « différence » humaine, de l'unicité des personnes. — L'individualiste aime cette « différence » non seulement en soi, mais chez autrui. Il est porté à la reconnaître, à en tenir compte et à s'y complaire. Cela suppose une intelligence fine et nuancée. Pascal a dit : « A mesure qu'on a plus d'esprit, on trouve qu'il y a plus d'hommes originaux. Les gens du commun ne trouvent pas de différence entre les hommes. » La sensibilité sociale ou grégaire se complaît dans la banalité des traits : elle aime qu'on soit « comme tout le monde ». La sensibilité chrétienne, humanitaire, solidariste et démocratique voudrait effacer les distinctions entre les moi. Amiel y voit avec raison l'indice d'une intellectualité grossière. « Si, comme dit Pascal, à mesure qu'on est plus développé, on trouve plus de différence entre les hommes, on ne peut dire que l'instinct démocratique développe beaucoup l'esprit, puisqu'il fait croire à l'égalité des mérites en vertu de la similitude des prétentions. » Le chrétien dit : « Faites à autrui ce que vous voudriez qu'il vous fît. » A quoi un dramaturge moraliste, B. Shaw, réplique avec esprit : « Ne faites pas à autrui ce que vous voudriez qu'il vous fît ; vous n'avez peut-être pas les mêmes goûts [1]. »

1. G. Palante, *Mercure de France* du 16 juin 1908 ; les caractères de la *sensibilité individualiste* sont définis dans cette étude avec beaucoup de finesse ; mais plusieurs de ces caractères sont, quoi que paraisse en penser l'auteur, communs à presque tous les hommes ; c'est une question de degré.

Ce besoin d' « unicité » ne se manifeste pas d'ailleurs seulement chez quelques rares individus privilégiés ; il existe chez presque tous les hommes, prenant parfois des formes naïves et un peu ridicules, dont l'une des plus curieuses est l'importance que certaines personnes attachent à tout ce qui concerne leur propre nom. L'importance attachée à la reproduction du nom prend les formes les plus diverses : la plus grossière est peut-être le désir de voir son nom écrit partout, désir dont l'expression, suivant les classes sociales, évolue des « graffitti » aux « échos mondains » ; une minime erreur d'orthographe dans une adresse de lettre, le fait de n'être pas appelé immédiatement par son nom, même si on est reconnu, provoquent parfois une réelle mauvaise humeur. Ces diverses formes de susceptibilité sont d'ailleurs particulièrement répandues chez les esprits les plus médiocres, ayant le moins d'individualité : il est, somme toute, assez naturel qu'ils cherchent à accroître par tous les moyens leur inconsistante personnalité.

Un fait plus normal est que l'homme n'aime pas, en général, perdre son nom et être désigné par un numéro ; ni même être compté seulement comme une unité dans un groupe sans être individuellement désigné. C'est déjà là une raison pour que la statistique ne soit pas populaire et pour que les plaisanteries faciles que l'on peut faire à son sujet soient généralement bien accueillies ; l'une des plus classiques consiste à énoncer gravement que telle voiture publique transporte journellement 245 voyageurs et 47 centièmes de voyageur ; l'élément comique résulte précisément de ce que l'on pousse jusqu'à l'absurde l'assimila-

tion de l'individu avec une simple unité arithmétique abstraite.

Voilà donc un premier élément de suspicion contre la théorie des probabilités dans tous les cas fort nombreux où elle se base sur le dénombrement d'individus humains; mais ce n'est point le seul, ni le plus important.

La théorie des probabilités ne se confond pas en effet avec la statistique, à laquelle s'applique aussi tout ce qui précède ; non contente de recenser les événements passés, elle prétend prévoir dans une certaine mesure les événements futurs : c'est en cela qu'elle est une science. Cette prétention heurte tout d'abord le sentiment psychologique de la liberté humaine (dont la valeur métaphysique n'est pas ici en question) ; l'on affirme que, s'il ne se produit pas d'événements exceptionnels, tels que guerre, tremblement de terre, etc., il y aura sûrement plus de 1.000 mariages en France la semaine prochaine : ne dépend-il pas des fiancés de démentir la prédiction en ajournant d'une semaine la célébration de leur union[1] ? L'objection ne supporte pas l'examen, mais elle est souvent admise implicitement sans examen.

1. Il semble que Joseph Bertrand ait pensé, sans toutefois jamais le dire expressément, que la clause « s'il ne se produit pas d'événement exceptionnel » enlève toute valeur aux prédictions basées sur le calcul des probabilités. Cette clause a pourtant un sens très clair et intervient dans tous les phénomènes scientifiquement observables : l'astronome verra passer telle étoile à telle heure au méridien, *si sa lunette et son horloge fonctionnent bien.* Particulièrement dans les événements où intervient la liberté humaine, il est sans intérêt de constater qu'un motif exceptionnel peut causer des perturbations ; par exemple une épidémie provoque un exode exceptionnel de voyageurs : chacun a alors conscience que sa décision a été influencée par ce motif : ce qui est intéressant, c'est de constater la régularité des résultats lorsque les mobiles sont variés suivant les individus.

92. — Il est enfin des cas où les prévisions du calcul concernent des phénomènes indépendants de la volonté de l'homme, mais qui le touchent de trop près pour qu'il puisse les considérer froidement, aux seules lumières de la raison. Je ne parlerai pas des superstitions et des folies auxquelles la passion du jeu conduit les esprits les mieux équilibrés ; mais je voudrais dire quelques mots de l'application du calcul aux probabilités concernant la vie humaine. Il est peu de sciences dont la valeur objective soit moins contestable ; les compagnies d'assurances sur la vie, du moins celles dont la gestion est sérieuse et conforme aux règles scientifiques, encaissent des bénéfices et distribuent des dividendes dont la réalité tangible ne redoute aucun scepticisme. Mais, si l'on ne peut contester ces résultats généraux, certaines affirmations particulières, souvent mal comprises, se heurtent à l'appréhension superstitieuse ressentie par beaucoup d'hommes dès que, même très indirectement, il peut être question de leur propre mort. Cet « individualisme biologique » est très répandu et a des manifestations très diverses, dont la plus souvent signalée est un certain besoin puéril d'affirmer l'excellence de sa propre santé, la « solidité de son coffre ». Chacun jugera volontiers les prévisions du calcul fort raisonnables, à condition qu'on les applique à une population dont il ne fait pas partie, aux Chinois ou aux nègres s'il est de race blanche ; mais il n'aime point qu'on assigne la durée de sa vie moyenne ou probable, distinguant d'ailleurs mal le sens de ces expressions

Il est, en ces questions, une autre source de

paradoxe, qui apparaîtra clairement dans un exemple fourni par Joseph Bertrand[1].

Dans un problème plus célèbre et plus grave, la vie humaine servait d'enjeu. L'inoculation, avant la vaccine, était, contre la variole, le meilleur parti qu'on pût prendre; mais 1 inoculé sur 200 mourait des suites de l'opération. Quelques-uns hésitaient; Daniel Bernoulli, géomètre impassible, calculait doctement la vie moyenne, la trouvait accrue de trois ans et déclarait par syllogisme l'inoculation bienfaisante. D'Alembert, toujours hostile à la théorie du jeu, qu'il n'a jamais comprise, repoussait avec grande raison cette fois l'application qu'on en voulait faire : « Je suppose, dit-il, que la vie moyenne d'un homme de trente ans soit trente autres années et qu'il puisse raisonnablement espérer de vivre encore trente ans en s'abandonnant à la nature et en ne se faisant pas inoculer. Je suppose ensuite qu'en se soumettant à l'opération la vie moyenne soit de trente-quatre ans. Ne semble-t-il pas que, pour apprécier l'avantage de l'inoculation, il ne suffit pas de comparer la vie moyenne de trente-quatre ans à la vie moyenne de trente, mais le risque de 1 sur 200 auquel on s'expose, de mourir dans un mois, par l'inoculation, à l'avantage éloigné de vivre quatre ans de plus au bout de soixante ans? »

On argumente mal pour vider de telles questions : supposons que l'on puisse, par une opération, accroître la vie moyenne, non plus de quatre, mais de quarante ans, à la condition qu'une mort immédiate menacera le quart des opérés : un quart des vies sacrifié pour doubler les trois autres, le bénéfice est grand. Qui voudra le recueillir? Quel médecin fera l'opération? Qui se chargera, en y invitant 4.000 habitants robustes et bien portants d'une même commune, de commander pour le lendemain 1.000 cercueils? Quel directeur de collège oserait annoncer à 50 mères, qu'empressé à accroître la vie moyenne de ses 200 élèves, il a joué pour eux ce jeu avantageux et que leurs fils sont les perdants? Les parents les plus sages acceptaient une chance sur 200; aucun, sur la foi d'aucun calcul, ne s'exposerait à 1 chance sur 4.

1. *Calcul des probabilités*, p. XII.

Le paradoxe est clairement exposé, mais ni d'Alembert, ni Bertrand ne mettent en évidence la raison profonde pour laquelle « on argumente mal » en de telles questions. Cette raison me paraît être la suivante : l'incertitude de la date de sa mort est une condition essentielle de la vie de l'homme ; il est aussi absurde de renoncer actuellement à cette donnée que de supposer des incursions sur les planètes voisines. Si, dans l'avenir, les progrès de la science permettaient d'assigner la date exacte de la mort de chacun, la mentalité de l'humanité en serait modifiée au point que nous ne pouvons prévoir quelle serait son attitude en face des problèmes que pose Bertrand. Et, cependant, cette hypothèse est nécessaire pour que ces problèmes puissent être réellement posés d'une manière concrète : sinon, ce ne sont que des spéculations vides de sens. Le jour où l'on pourrait dire à un homme de trente ans : « vous êtes scientifiquement assuré de vivre jusqu'à soixante ans, mais pas davantage ; telle opération prolongera *sûrement* votre vie jusqu'à quatre-vingts ans, mais vous risquez une mort immédiate, qui se produit une fois sur dix », et où la confiance dans une science éprouvée ne laisserait aucun doute sur la vérité de ces affirmations, ce jour-là on pourrait se demander quelle réponse doit faire cet homme. Mais, pour nous qui vivons actuellement, de telles affirmations seraient sans fondement ; nous savons très bien qu'on ne peut nous assurer l'avenir et cette incertitude même est nécessaire pour que nous goûtions la vie.

On pourra répondre que le calcul des probabilités ne saurait promettre une certitude, mais

peut garantir, dans certains cas, une moyenne ; n'est-ce pas un résultat positif que d'augmenter cette moyenne ? Sans doute, mais alors il ne faudrait pas raisonner comme si la vie réelle devait être toujours exactement égale à cette moyenne ; la répartition autour de la moyenne n'est nullement indifférente, et c'est ce qu'oublient d'Alembert et Bertrand. Etudions de plus près un de leurs exemples.

Bertrand considère 100 individus, dont chacun doit vivre quarante ans et propose d'en sacrifier immédiatement 25 pour prolonger pendant quatre-vingts ans la vie des 75 autres ; il n'a pas de peine à montrer que cette proposition est inacceptable. Il n'en serait pas forcément de même si l'on admettait, comme on doit le faire, que chacun ne vit pas précisément la vie moyenne. Supposons 100 individus atteints d'une maladie grave, dont on doit supposer que 50 mourront dans un très bref délai ; les 50 autres vivront *en moyenne* quarante ans ; la durée moyenne de la vie, *pour les* 100 *individus*, est donc vingt ans environ ; si, par telle opération ou tel traitement, on en sauve 75 dont la vie moyenne sera encore quarante ans, les 25 autres succombant à l'opération, la durée moyenne de la vie, *pour les* 100 *individus*, deviendra trente ans ; ici, il n'est pas douteux que l'augmentation de la vie moyenne ne soit avantageuse, et un médecin bien renseigné conseillera le traitement qui sur 4 malades en sauve 3 au lieu de 2 seulement.

Supposons, au contraire, que par une opération sur les nouveau-nés, on évite telle maladie de la vieillesse et qu'on prolonge ainsi de quelques années la vie moyenne de ceux qui dépasseront

soixante ans ; même si le calcul prouve que la durée moyenne de la vie est ainsi augmentée pour l'ensemble des individus, aucune mère ne soumettra son enfant à l'opération, si celle-ci est mortelle seulement 1 fois sur 100. Tout cela ne prouve qu'une chose : c'est que, pour interpréter les résultats d'un calcul, on n'est pas dispensé d'avoir du bon sens.

93. — Mais le bon sens, non plus que le calcul, n'assure contre le malheur, et ce sera toujours une maigre consolation pour un individu de penser que la probabilité du malheur était faible, si c'est lui qui le subit. Celui qui meurt de faim s'intéresse peu à l'augmentation de la fortune moyenne : on ne doit pas chercher dans la statistique ni dans le calcul des arguments pour consoler ceux qui souffrent des inégalités sociales ; cette constatation ne diminue en rien la valeur propre des statistiques ni des calculs par lesquels on les interprète.

C'est seulement, en effet, à un point de vue particulier que le statisticien ou le mathématicien étudient les phénomènes sociaux ; cette étude a une portée limitée, mais elle constitue une science exacte lorsque l'on ne prétend pas l'étendre au delà de ses limites naturelles. On ne doit y chercher ni arguments moraux, ni raisons immédiates d'agir : mais seulement, comme dans les sciences physiques, un moyen de bien connaître les événements passés et de prévoir avec une certaine approximation les événements futurs. Lorsque l'on prédit que plus de cent mille Parisiens prendront demain le « métro », on n'oblige aucun d'eux à choisir ce moyen de transport ; on

énonce simplement un fait que l'expérience confirme. De même, si l'on prédit que, telle municipalité annonçant l'entreprise de plusieurs millions de travaux, on peut en conclure que quelques ouvriers périront victimes d'accidents du travail, on ne mérite, pas plus que ceux dont l'initiative ordonne les travaux, d'être traité d'assassin : rien ne prouve que la mortalité n'aurait pas été plus grande chez les ouvriers inoccupés, ou occupés ailleurs ; substituer à cette incertitude une prévision relativement précise augmente nos connaissances sans faire de tort à personne.

Il faut seulement prendre garde de ne pas tomber dans un travers trop répandu : bien des personnes redoutent d'autant plus les inconvénients et les dangers qu'ils sont mieux connus ; elles seront effrayées par une statistique précise des accidents de chemins de fer et ne penseront pas aux accidents moins soigneusement relevés atteignant les personnes qui se promènent tranquillement à pied dans les rues ou sur les grandes routes. De même, leur attention ayant été attirée sur les inconvénients précis que peut avoir tel mode d'alimentation, elles s'imposeront un régime bizarre, sans se demander si les modifications inconnues produites dans leur organisme par ce régime anormal ne les exposent pas à des dangers plus graves que ceux qu'elles cherchent à éviter.

L'ignorance peut être commode pour ceux qui pratiquent cette politique d'autruche ; elle n'est jamais désirable pour ceux qui préfèrent voir clair et ne se laissent pas influencer par la connaissance plus exacte qu'elles acquièrent d'un danger possible, lorsque sa probabilité est nota-

blement inférieure à celle des dangers inconnus auxquels les hommes les plus timorés s'exposent tous les jours[1]. On n'a rien à redouter du calcul, lorsqu'on est décidé à ne pas régler sa conduite sur ses indications sans les avoir au préalable pesées à leur juste valeur : c'est une illusion singulière que de penser que l'indépendance individuelle est accrue par l'ignorance.

94. — N'y a-t-il donc aucun fondement à l'opposition que nous avions cru discerner entre la théorie des probabilités et l'individualisme ? Il en est un très réel au contraire, dans la mesure où l'individualisme est antisocial, car la théorie des probabilités est la base de ce que l'on peut appeler les *mathématiques sociales*. Son étude nous rappelle en effet que nous vivons en société et que les phénomènes sociaux ont une existence réelle et un intérêt propre. Elle nous rappelle que, si les hommes sont différents en bien des points, ils sont cependant semblables en ce qu'ils sont tous exposés aux accidents, à la maladie, à la mort; en ce que leurs divers éléments biologiques (la taille, les dimensions du crâne, etc.), se répartissent suivant des lois régulières autour de certaines moyennes, en ce qu'enfin on peut énoncer des *lois* que vérifie l'observation des faits et dans l'énoncé desquelles ils sont considérés comme des unités constitutives de l'ensemble auquel s'applique la loi. De telles constatations sont éminemment propres à limiter les excès de l'égoïsme individualiste. Telle maladie fait, en moyenne, un certain nombre de victimes;

[1] Voir plus haut, n° 88.

la grêle ou les inondations font, en moyenne, un certain chiffre de dégâts ; on ne sait pourquoi les uns sont atteints et les autres épargnés, mais la société, dans son ensemble, subit un dommage à peu près constant. L'étude de ces faits ne peut que contribuer à développer la notion de la solidarité, à rappeler à chacun qu'il ne doit pas se considérer comme indépendant du milieu où il vit et qu'il doit participer à la réparation des dommages fortuits qui atteignent son voisin et auraient pu l'atteindre lui-même. Aussi l'étude de la théorie des probabilités a-t-elle une très grande valeur éducative ; on devrait souhaiter qu'elle pût être mise à la portée de tous ceux qui prétendent à une part dans la direction des hommes et des choses. Ainsi serait fort heureusement combattu un certain individualisme qui n'est autre chose qu'un égoïsme inintelligent. Mais la théorie des probabilités ne menace nullement le véritable individualisme, c'est-à-dire la conscience nette de l'indépendance de pensée et d'action d'une personnalité qui se sent libre.

95. — Est-il possible d'aller plus loin, et de fonder sur la théorie des probabilités une véritable morale individuelle et sociale? C'est là un problème que je ne puis avoir la prétention de traiter ici, mais je voudrais indiquer tout au moins dans quelle voie on pourrait en chercher la solution.

Il semble que la règle de morale la plus élevée qui ait été proposée aux hommes se résume dans la parole évangélique « Tu aimeras ton prochain comme toi-même ». L'illustration pratique la plus populaire en est peut-être l'histoire de saint

Martin partageant son manteau. Si l'on ose appliquer à un tel sujet la critique scientifique, au risque d'être accusé d'une excessive sécheresse de cœur, on est bien obligé de reconnaître que si le prochain, ce sont *tous* les hommes, l'application stricte du précepte évangélique conduit à des conséquences absurdes. Saint Martin lui-même n'aurait pu partager son manteau en un millier de morceaux s'il eût rencontré un millier de malheureux. On peut mettre cette remarque banale sous une forme arithmétique en observant que l'homme qui regarderait tous les habitants du globe, ou même de sa patrie, ou même de sa cité, pour peu que ce soit une grande ville, comme équivalents à lui-même, serait conduit à partager non seulement ses biens, mais encore son activité, en un si grand nombre de parts que sa vie serait rendue impossible. On peut objecter que si chacun l'imitait, il y aurait dans ce communisme total une compensation qui permettrait à tous de vivre, mais la dispersion indéfinie des activités détruirait la personnalité même des individus. La seule interprétation raisonnable qui puisse être donnée de la maxime évangélique est donc la suivante : considère chacun de tes prochains, non comme équivalent en tout cas à toi-même, mais comme équivalent à une fraction de toi-même comprise entre *zéro* et *un*, mais n'atteignant jamais la limite inférieure *zéro*, tandis qu'elle peut parfois atteindre la limite supérieure *un*. Je ne pense pas qu'une telle formule puisse être taxée d'égoïste ; du moment que l'on en a bien compris la signification, elle est au contraire la formule la plus large et la plus complète de l'altruisme intelligent. Les degrés divers

dans l'altruisme et dans l'égoïsme se révéleront à la fixation des coefficients : à combien de personnes chacun de nous attribue-t-il le coefficient 1, à combien le coefficient 0,9, le coefficient 0,5, ... le coefficient 0,000001 ? Je me garderai d'entrer dans la discussion de ces valeurs, qui est du ressort de la morale pratique[1] ; le point essentiel, c'est que ces coefficients ne doivent pas être nuls, et l'on peut regarder cette affirmation comme le fondement de la morale théorique. Dans la constatation de ce fait et dans l'étude de ses conséquences, la théorie des probabilités intervient à chaque instant et donne une base objective aux aspirations du poète :

> O insensé, qui crois que je ne suis pas toi !

Il y a une probabilité non nulle pour qu'un tort fait à autrui nous soit dommageable, pour qu'un bénéfice réalisé par autrui nous soit profitable, pour que la maladie ou la mort de notre prochain nous soient fâcheuses, etc. D'autre part, les personnes les plus égoïstes évaluent plus ou moins consciemment le rapport entre le profit qu'elles espèrent d'un acte et le tort qu'il peut causer à autrui et n'accomplissent pas l'acte si ce rapport est trop faible ; mettre le feu à une meule de gerbes pour rôtir une alouette est regardé comme le fait d'un aliéné.

Il serait désirable que la relation entre le profit espéré et le tort causé, qui intervient dans nos jugements sur les actes, intervînt aussi dans la

1. Bien entendu, l'emploi de ces coefficients arithmétiques est surtout ici une forme abrégée de langage. La signification ne pourrait en être précisée que par une assez longue discussion dans laquelle on aurait le choix entre plusieurs conventions.

législation, dans la jurisprudence et dans les mœurs.

Le commerçant en pharmacie qui, grâce à une savante publicité, empêche par ses produits médiocres ou frelatés la guérison de milliers de malheureux sur chacun desquels il gagne quelques francs devrait apparaître comme bien plus coupable que l'assassin qui ne tue que deux ou trois personnes pour s'assurer la fortune.

96. — Une objection s'est certainement déjà présentée à l'esprit du lecteur : n'y a-t-il pas des cas où le coefficient attribué par tel homme à son prochain est non seulement nul, mais négatif ; en d'autres termes, ne peut-il arriver que l'on soit conduit à désirer le malheur du voisin, même si l'on n'en tire aucun profit direct et immédiat, et à plus forte raison si l'on a un profit, si léger soit-il. L'examen des diverses hypothèses pratiques sortirait de notre cadre ; je me bornerai à deux brèves observations.

L'homme qui cherche à tirer profit du malheur du voisin est un être anti-social contre lequel il est légitime que la société se défende ; il n'est pas nécessaire d'insister sur ce point. La peur du gendarme doit, en certains cas, suppléer à l'insuffisance de la conscience morale.

Mais la société n'est-elle pas organisée de telle manière que, dans certaines circonstances, le malheur des uns soit indirectement profitable à d'autres, sans que ceux-ci soient coupables d'aucune action répréhensible, la destruction de richesses entraînant, par exemple, la hausse des prix ? S'il en est ainsi, la société doit tendre à supprimer ces circonstances ; c'est là une règle

de morale sociale qui complète les règles de morale individuelle.

97. — Observons, en terminant, que l'on peut déduire des coefficients par lesquels peut être mesuré l'amour du prochain à la fois le patriotisme et l'internationalisme au sens le plus élevé. Pour chacun de nous, la somme des coefficients qu'il attribue à ses compatriotes est supérieure à la somme des coefficients attribués aux étrangers et, de plus, est très notablement supérieure à l'unité. Si donc un danger menace tous nos compatriotes, notre devoir est de les défendre en nous sacrifiant au besoin nous-même ; mais nous devons d'autre part désirer qu'il s'établisse entre les nations une morale de solidarité analogue à celle qui existe entre les individus. On sait combien est grande déjà actuellement la solidarité économique entre les nations, au point que l'on a pu se demander si une guerre victorieuse ne serait pas nuisible au vainqueur lui-même. Ceci revient à dire que si une nation s'attribue à elle-même un coefficient égal à l'unité, elle doit attribuer aux autres nations des coefficients inférieurs à l'unité, mais qui ne doivent être ni nuls ni négatifs[1]. Dire que les coefficients des autres nations sont inférieurs à l'unité, c'est dire qu'il doit subsister un certain égoïsme national, aussi nécessaire à la vie des peuples que l'égoïsme individuel l'est à la vie des individus. Dire que les coefficients des autres nations ne doivent pas être négatifs, c'est souhaiter que l'état de guerre ne soit pas substitué à l'état de paix.

1. Voir Norman Angell, *La Grande illusion*, et N. Murray Butler, *L'Esprit international*.

CHAPITRE IX

LA VALEUR SCIENTIFIQUE DES LOIS DU HASARD

98. Les lois du hasard et les diverses sciences. — 99. La méthode des majorités. — 100. Le calcul des probabilités et les décisions judiciaires. — 101, 102. La vérité absolue et la vérité relative. — 103. Cas où la « vérité relative » est sans intérêt. — 104. Recherche de la signification de la vérité relative. — 105. La méthode des « cas vrais et des cas faux ». — 106. Etude expérimentale des physionomies d'enfant. — 107. Etudes de Binet sur les mains. — 108. Conclusion.

98. — Il peut paraître superflu de parler de la valeur scientifique de la théorie des probabilités, après ce que nous avons dit de l'usage qui est fait de cette théorie dans des questions scientifiques diverses. Pour le savant, la science porte sa justification en elle-même : si la mécanique statistique conduit à des résultats importants, elle acquiert par là même une valeur scientifique aux yeux du physicien. Quant à la valeur philosophique de ces méthodes, nous la discuterons dans le chapitre suivant. Mais auparavant, il peut y avoir quelque intérêt à rechercher sous quelle forme la théorie des probabilités peut être appliquée à des questions pour lesquelles il est douteux si elles doivent être regardées comme proprement scientifiques ; en d'autres termes, nous allons rechercher si la théorie du hasard peut être

utilisée pour étendre les frontières même de la science[1].

99. — Nous donnerons d'une manière générale, le nom de *méthode des majorités* au procédé qui consiste à regarder comme pratiquement valable l'avis exprimé par la majorité, après dépouillement des avis ou des opinions d'un nombre plus ou moins grand de personnes. Nous ne nous attarderons pas, pour l'instant, à énumérer et discuter les formes diverses que peut prendre la méthode, ni à essayer de classer ses nombreuses applications, réalisées ou possibles. La question qui nous occupera principalement est la suivante : la connaissance du calcul des probabilités peut-elle être de quelque utilité à ceux qui emploient ou qui critiquent la méthode des majorités?

100. — Joseph Bertrand a rassemblé dans son *Calcul des Probabilités* (XLIII à XLIX et 319 à 327) les arguments qui opposent, pour ainsi dire, la question préalable à toute introduction du calcul des probabilités en ces matières. Quelques citations feront comprendre le point de vue auquel il se place.

L'application du calcul aux décisions judiciaires est, dit Stuart Mill, le scandale des Mathématiques. L'accusation est injuste. On peut peser du cuivre et le donner pour or, la balance reste sans reproche. Dans leurs travaux sur la

1. Les pages qui suivent jusqu'au nº 107 inclus ont été écrites en 1908 sur la demande du regretté Alfred BINET, pour son *Année psychologique* (t. XIV). Il m'a paru désirable, en raison des discussions des méthodes de Binet qui y figurent, de les reproduire ici sans modification.

théorie des jugements, Condorcet, Laplace et Poisson n'ont pesé que du cuivre.

... « Condorcet a pris possession de l'univers moral pour le soumettre au calcul. » C'est la louange qu'on lui a donnée; on s'est demandé si c'est après l'avoir lu. Dans son livre sur *la probabilité des jugements*, il se propose d'abord deux problèmes. Premièrement : Quel est, pour chaque jugement et pour chaque juge, la probabilité de rencontrer juste ? En second lieu, quelle est la probabilité d'erreur à laquelle la société peut se résigner sans alarmes ?

La première question lui semble facile.

« Je suppose, dit Condorcet, que l'on ait choisi un grand nombre d'hommes véritablement éclairés et qu'ils prononcent sur la vérité ou la fausseté de la décision. Si, parmi les décisions de ce tribunal d'examen, on n'a égard qu'à celles qui ont obtenu une certaine pluralité, il est aisé de voir qu'on peut, sans erreur sensible, les regarder comme certaines. »

C'est un concile infaillible, tout simplement, qu'il définit et prétend convoquer. Sans douter, il hésite; non que les hommes véritablement éclairés soient rares, gardons-nous de le croire, mais leur temps est précieux. Pour l'épargner, Condorcet propose une seconde méthode dont Poisson, plus tard, n'a pas aperçu l'illusion. La probabilité d'erreur étant supposée connue pour un juré, on peut, sans augmenter leur nombre, la diminuer sans limite pour l'ensemble.

... Si, comme le demande très sérieusement Cournot, on invitait le greffier à noter, après chaque jugement, l'opinion de chacun des juges, pour appliquer, quand les chiffres seront nombreux, la formule qui donne leur mérite, la perspicacité de chacun étant contrôlée par celles de ses deux collègues, le juge le mieux noté de France serait celui qui, sans discuter ni réfléchir, voterait toujours comme son président : s'il faut en croire la formule, un tel juge ne se trompe jamais.

Ni Cournot ni Poisson n'ont commis la plus petite faute comme géomètres; ils traduisent rigoureusement leurs hypothèses. Mais les hypothèses n'ont pas le moindre rapport avec la situation d'un accusé devant les juges.

Ils ont aperçu les différences et croient en les signalant, acquérir le droit de n'en pas tenir compte.

Poisson qui, comme Condorcet, a consacré à la théorie des jugements un volume entier rempli des plus savants calculs, croit atténuer les objections qu'il ne pouvait manquer d'apercevoir, en altérant, dans ses énoncés, la signification du mot *coupable*. On rendrait, dit-il, le langage plus exact en substituant le mot *condamnable*, qui est toute la vérité, au mot *coupable*, qui avait besoin d'explications et que nous continuerons d'employer pour nous conformer à l'usage.

L'innocent accablé sous des indices trompeurs ou victime de machinations trop habiles pour qu'aucun juge puisse les soupçonner est un accusé *condamnable*. Poisson, *pour se conformer à l'usage*, le classe parmi les coupables. L'erreur unanime des juges devient alors une preuve de sagacité dont l'algèbre leur tient compte en évaluant leur mérite avec une infaillible précision. Dans cette suite de calculs stériles qui resteront, comme l'a dit Stuart Mill, le scandale des mathématiques, Condorcet seul a donné un sage conseil : celui de choisir, pour composer les assemblées, des hommes *véritablement éclairés*.

J'ai tenu à faire ces citations, où l'ironie et le paradoxe cachent le plus souvent une pensée très juste ; il y a beaucoup à retenir des appréciations de Bertrand, et il était utile de les rappeler. Il serait intéressant de rechercher si Condorcet, Laplace, Poisson, Cournot furent vraiment aussi naïfs que Bertrand semble le croire ; mais je n'entrerai pas dans cette discussion historique, assez inutile pour notre but ; il ne s'agit pas de savoir si les affirmations de Condorcet sont justes, ou si Bertrand a raison de les critiquer ; mais bien de déterminer sous quelle forme on peut actuellement songer à appliquer le calcul des probabilités à la méthode des majorités. Dans cette étude, les travaux antérieurs nous seront naturellement utiles ; mais les citer et les discuter à chaque instant ne pourrait qu'alourdir sans profit l'exposition.

101. — Le premier point qui est de toute évidence, c'est que la majorité ne peut pas posséder des lumières qui font complètement défaut à chacune des individualités qui la composent : un jury d'aveugles ne saurait juger des couleurs. L'impossibilité, qui tient ici à la nature du jury, peut, dans d'autres cas, être inhérente à la question posée : *connaissant la hauteur du grand mât, trouver l'âge du capitaine;* mille personnes ne résoudront pas cette question mieux qu'une seule.

Ces remarques paraîtront peut-être superflues, tellement elles sont évidentes; on peut cependant en tirer une conséquence qui est fondamentale dans notre étude : *la valeur du jugement de la majorité peut dans certains cas être plus grande que la valeur du jugement individuel, mais* ELLE NE PEUT PAS ÊTRE DE QUALITÉ DIFFÉRENTE.

Un exemple très simple fera bien comprendre le sens de cet énoncé : Paul joue à l'écarté, et est très inexpérimenté; il se trouvé avoir la dame, l'as et le neuf de cœur, qui est atout, le roi de trèfle et le roi de pique. Il se demande s'il doit jouer d'autorité, et consulte des amis plus éclairés : ceux-ci l'y encouragent, mais son adversaire a le roi, le valet et le dix d'atout et deux petits carreaux; Paul perd la partie. A-t-il eu tort de suivre les conseils qu'on lui a donnés? Ou, plus précisément, doit-on dire que ces conseils étaient mauvais? Assurément non; ces conseils étaient aussi bons que possible, du moment que les conseilleurs ne connaissaient pas le jeu de l'adversaire de Paul : seul, un compère habile et peu honnête aurait pu, ayant entrevu ce jeu, donner à

Paul un conseil destiné à mieux réussir dans la circonstance très particulière où il se trouvait.

Nous pouvons résumer ceci en disant que les conseils donnés à Paul par ses amis constituent la *vérité relative*, dans les conditions où se trouve Paul, c'est-à-dire dans l'ignorance où il se trouve du jeu de son adversaire ; mais il peut arriver que cette *vérité relative* soit contradictoire à la *vérité absolue*, c'est-à-dire à la manière de jouer qui serait la meilleure pour celui qui connaîtrait les deux jeux. Il est d'ailleurs bon d'ajouter que cette vérité, que nous appelons *absolue*, est seulement *moins relative;* il peut se faire que la perte de cette partie, influant sur la disposition des cartes pour la partie suivante, soit, en fin de compte, avantageuse à Paul, de sorte que le conseilleur qui connaîtrait, non seulement le jeu de son adversaire, mais la manière dont les cartes seront ramassées et battues pour le coup suivant, pourrait donner un conseil différent et pratiquement meilleur. Il est donc bien entendu que le mot *absolu* ne signifie que *moins relatif;* il ne peut d'ailleurs avoir d'autre sens, dans toute étude portant sur des réalités.

En admettant cette manière de parler, on peut distinguer, au point de vue pratique, trois catégories principales dans les applications du calcul des probabilités à la méthode des majorités :

1° La *vérité relative* que l'on atteint a une signification intéressante par elle-même, au point qu'il est légitime de la prendre comme but absolu de la recherche ;

2° La vérité relative que l'on atteint n'a qu'un rapport éloigné et inconnu avec la vérité absolue qui seule intéresse ;

3° La question est posée de telle manière que l'on doute s'il y a une vérité relative; on doit donc se demander si cette vérité relative existe; on pourra ensuite se proposer d'en déterminer la nature.

Ces distinctions, je le répète, ne peuvent avoir qu'une valeur *pratique;* au point de vue théorique, tous les cas intermédiaires sont possibles; au point de vue abstrait, tous les cas sont semblables en ce que leur traduction mathématique est la même; mais la signification utile de ces mêmes formules est fort différente suivant les cas : la même balance, dirait Joseph Bertrand, pèse tantôt du cuivre et tantôt de l'or.

102. — Occupons-nous d'abord des cas où la vérité relative, non seulement est intéressante par elle-même, mais doit être regardée comme ayant une valeur absolue : ce sont les cas où le fait crée le droit.

L'un des meilleurs exemples que l'on puisse en donner est celui des questions de langage[1]; l'on se propose, par exemple, de savoir si telle locution est actuellement usitée ou comprise en un lieu donné : le moyen le plus sûr est d'interroger un grand nombre des habitants de ce lieu; si on pouvait les consulter tous, individuellement et indépendamment l'un de l'autre, l'ensemble de leurs réponses, supposées sincères et exactement dénombrées, fournirait évidemment, par définition même, la réponse la plus satisfaisante possible à la question posée. Mais il est clair qu'une

1. Voir plus haut n° 47.

telle consultation est pratiquement impossible ; la méthode des majorités consistera donc, en consultant un petit nombre seulement de personnes, à chercher à prévoir le résultat que donnerait le dénombrement général. Par exemple, pour savoir si les Parisiens de 1908 disent *un* automobile ou *une* automobile, on consultera vingt personnes, et on prendra l'avis de la majorité. C'est ce que font souvent les journaux lorsqu'ils instituent une « enquête ». Que vaut la méthode ? Sans entrer dans les détails, il est nécessaire, pour que le calcul puisse être correctement appliqué à une question, que le groupe total par rapport auquel cette question est posée, soit sensiblement homogène ; cette homogénéité n'est jamais absolue, puisque deux individus quelconques ne sont pas identiques ; il faut donc de plus que le groupe partiel que l'on choisit comme représentant du groupe total ait sensiblement la même hétérogénéité. Par exemple, il serait absolument incorrect de prendre les cinquante voyageurs qui sont dans un tramway déterminé comme représentants de l'ensemble des Parisiens ; il serait plus correct de prendre comme représentants de l'ensemble des conscrits parisiens nés en 1887, ceux d'entre eux qui sont nés, par exemple, le 14 mai de cette année ; ce choix serait meilleur que le choix d'un quartier déterminé, ou même qu'un choix alphabétique ; en supposant les noms inscrits par lettre alphabétique et choisissant une portion de la liste désignée par le sort, on risquerait de tomber sur des portions trop homogènes, les personnes dont les noms commencent par *La* ou *Le*, ou bien par W ou Z pouvant avoir en commun certains

caractères ethniques les différenciant des autres.

Si les conditions d'homogénéité sont réalisées, c'est-à-dire si le groupe restreint sur lequel porte l'expérience est vraiment une image fidèle du groupe total, quelles conclusions peut-on tirer de l'étude de ce groupe partiel ? Par exemple, sur 36.500 conscrits d'une région donnée nés en 1887, il y en a 100 qui sont nés le 14 mai ; on constate que, sur ces 100, la majorité absolue, soit 51, a une taille supérieure ou égale à 1 m, 65, les 49 autres ayant une taille inférieure ; on constate, d'autre part, sur la majorité des 100, telle ignorance, ou telle opinion, ou tel jugement particulier ; que peut-on induire de là pour l'ensemble total dont ils ont été extraits ? Ce n'est point ici le lieu de traiter mathématiquement ces questions ; il suffit d'avoir indiqué qu'il est légitime de se les poser et de signaler la forme de la réponse qu'y fournit le calcul. Cette forme est nécessairement un coefficient de probabilité. Par exemple, on pourra, des faits observés, conclure ceci : la probabilité pour que la majorité des faits non observés soit conforme à la majorité des faits observés est de 0,999, la probabilité opposée étant par suite 0,001. En d'autres termes, on peut parier 999 contre 1 que l'on ne se trompe pas, mais on n'a pas la certitude de ne pas se tromper. Cette forme particulière d'affirmation est commune à toutes les questions où intervient la théorie des probabilités ; certains esprits se refusent à la comprendre et préfèrent déclarer qu'ils veulent ignorer une théorie qui conduit à des résultats aussi incertains ; ils ne se rendent pas compte que cette incertitude est commune à toutes nos affirmations, et n'est pas moins dangereuse lorsqu'elle est

masquée sous des apparences dogmatiques[1]. En réalité, au point de vue pratique, une probabilité suffisamment voisine de l'unité doit, dans l'action, être confondue avec la certitude.

Indiquons enfin que la théorie des probabilités permet, de l'étude des chiffres observés, de tirer des conclusions sur l'homogénéité du groupe étudié et, par suite, sur la valeur que l'on peut attribuer aux observations faites sur ce groupe. Je suppose, par exemple, que les 10 premiers conscrits examinés au conseil de révision de la Seine aient mesuré, cinq d'entre eux 1 m. 55 et cinq d'entre eux 1 m. 75 ; la simple constatation du fait qu'aucun d'eux n'a une taille voisine de la moyenne 1 m. 65 suffit pour inspirer les doutes les plus sérieux sur l'exactitude des conclusions que l'on obtiendrait en supposant que ce groupe partiel est l'image exacte du groupe total dont il est extrait : on doit, au contraire, être certain que ces dix individus ne sont pas représentatifs du groupe.

103. — Je serai très bref sur le cas où la *vérité relative*, que la méthode des majorités permet d'atteindre, est sans rapport réel, ou du moins sans rapport connu, avec la vérité absolue qu'il serait utile de connaître. Ces cas ne sont pas intéressants : on doit les signaler seulement, à titre d'indication, afin d'éviter les erreurs auxquelles leur étude pourrait conduire. On peut se demander si toutes les applications du calcul aux décisions judiciaires rentrent dans cette catégorie, c'est-à-dire si la distinction faite par Poisson

1. Voir plus haut nos 85 à 89.

entre *coupable* et *condamnable* suffit pour justifier les sévérités de Joseph Bertrand. Il semble que celui-ci ait posé la question sous une forme trop abstraite. Pour obtenir des jugements équitables, il est clair qu'il faut avant tout choisir des juges éclairés et dépourvus de passion : mais ce n'est pas ainsi que la question est posée dans la pratique, et il est vraiment trop simple de proposer cette solution. Il ne s'agit pas de réaliser un idéal impossible à atteindre, mais d'arriver au moindre mal en se servant des institutions imparfaites et des hommes faillibles dont on dispose, jusqu'au jour, que nous souhaitons tous avec Joseph Bertrand, mais sans y compter plus que lui, où les institutions seront parfaites et les hommes infaillibles.

Lorsque l'on quitte ainsi le terrain de la justice idéale pour se placer sur celui des faits, on doit admettre qu'il existe des circonstances où tel accusé innocent sera fatalement condamné par la quasi-unanimité des juges ou jurés appelés à prononcer sur son sort. Il faut naturellement déplorer qu'il en soit ainsi et tâcher, en améliorant les garanties que la procédure donne aux accusés, en améliorant surtout la valeur des juges, de rendre un tel cas de plus en plus improbable, mais le désir d'améliorer la réalité ne doit pas empêcher de la constater. Dès lors, il est parfaitement légitime de se poser une question telle que la suivante : en augmentant le nombre des jurés et en modifiant la majorité requise pour la condamnation, augmente-t-on ou diminue-t-on les chances pour qu'un accusé soit condamné, lorsque son innocence non seulement est réelle, mais encore n'est pas masquée par des appa-

rences mensongères et des machinations habiles.

Lorsque la question est ainsi posée, il ne semble pas douteux que les calculs de Laplace et de Poisson n'y fournissent la réponse correcte[1]. Seulement, pour que cette réponse reste correcte, il ne faut pas en modifier le sens en faisant abstraction des différences concrètes qui existent entre les divers cas. Par exemple, il résulte aisément d'un calcul simple qu'il est préférable pour un accusé d'être jugé par un tribunal de 7 juges, où 5 voix sont nécessaires pour la condamnation, que par un tribunal de 12 juges où 7 voix suffisent à condamner. Mais il serait évidemment abusif d'en conclure qu'il est toujours préférable pour un soldat d'être traduit devant un conseil de guerre que devant un jury civil. Car, indépendamment de toute autre considération, le seul fait que le vote du jury a lieu au scrutin secret est un élément dont il a été impossible de tenir compte dans les calculs.

Parmi les cas où il n'y a *rien* à trouver, et où par conséquent la méthode des majorités et toute méthode statistique perd ses droits, il n'est pas inutile de mentionner les jeux de hasard, sans trop compter toutefois être cru par les joueurs, gens à préjugés tenaces et plus disposés à écouter

1. Il est cependant un point que Bertrand paraît avoir négligé et dont l'importance peut parfois être grande : la connaissance des conditions requises pour la condamnation ne peut-elle influer sur l'opinion des juges ? Par exemple, parmi les conventionnels ayant voté la mort de Louis XVI, ne peut-on pas regarder comme très vraisemblable qu'il en est au moins un qui se fût prononcé pour une condamnation moins sévère si, juge unique ou membre d'un jury très peu nombreux, son vote lui eût paru plus décisif? Cette question est distincte de celle du vote secret des juges, qui fait partie des dispositions pouvant être modifiées par la loi et les mœurs.

les prometteurs de martingales que les conseils de l'algèbre. Il est fâcheux de voir des chercheurs dont quelques-uns sont sérieux perdre à des recherches aussi vaines un temps qui pourrait être mieux employé.

La question de savoir si l'on peut gagner de l'argent à coup sûr en jouant à la roulette est une question essentiellement pratique et concrète ; à ce titre, elle appartient entièrement à la science et les arguments métaphysiques ou subjectifs n'ont rien à voir avec la solution pratique de cette question, solution qui a été donnée depuis longtemps par Bernoulli et confirmée par tous les mathématiciens : un joueur qui joue indéfiniment à un jeu équitable arrive forcément à la ruine, au bout d'un temps plus ou moins long ; ce temps devient très court si le jeu n'est pas équitable, ce qui est toujours le cas dans la pratique.

104. — Nous arrivons maintenant au cas le plus intéressant pour les psychologues, car il se présente souvent dans leurs recherches, bien que la question ne soit pas toujours posée par eux sous la forme que nous allons lui donner. Un ensemble de phénomènes étant observés par des individus divers, puis classés de telle manière qu'une majorité s'en dégage, le premier problème qui se pose est le suivant :

L'opinion de cette majorité correspond-elle à quelque réalité ?

Déterminer quelle est cette réalité et si, en particulier, elle est adéquate ou non à ce que l'on aurait pu espérer, en instituant les expériences, c'est un second problème, dont je ne méconnais

nullement l'importance, mais qu'il est essentiel de distinguer du premier, car c'est seulement lorsque le premier sera résolu que le second pourra être abordé avec fruit.

Précisons cette distinction par un exemple ; supposons que l'on ait lu à haute voix une même page dans un grand nombre d'écoles primaires et demandé à chaque enfant d'indiquer par écrit quel est le mot qui revient le plus souvent dans cette page[1]. Si les trois quarts des réponses s'accordent pour désigner le même mot, on peut conclure avec certitude que cette coïncidence n'est pas fortuite, mais a une raison. Cette raison peut être le fait que le mot désigné est effectivement celui qui figure le plus souvent dans la page lue ; elle peut aussi être tout autre ; mais c'est l'existence même d'une raison qui est le résultat essentiel fourni par la méthode des majorités dans cet exemple particulier.

De même si, faisant comparer à de nombreux expérimentateurs les poids de deux objets A et B dont les formes et dimensions sont très différentes, on obtient des réponses dont les trois quarts s'accordent pour attribuer un poids plus élevé à l'objet A,. on peut en conclure qu'il y a une raison à cela, mais cette raison n'est peut-être pas que A soit effectivement plus lourd que B. Au contraire si, sur 1 000 expériences, 520 seulement déclarent A plus lourd et 480 B plus lourd (je suppose qu'on laisse de côté les réponses douteuses), l'écart observé est de ceux qui se produi-

1. Une expérience analogue a été faite récemment par M. Binet ; on lisait à des enfants 100 substantifs et on leur demandait d'écrire la liste des mots dont ils se souvenaient. Je n'ai pas eu les éléments suffisants pour soumettre les résultats au calcul.

raient fréquemment dans une série de 1.000 réponses tirées à pile ou face ; on ne doit donc rien en conclure, sinon une forte présomption pour qu'une nouvelle série de 1.000 expériences fournisse un résultat aussi incertain.

Avant d'arriver à la discussion d'expériences à résultats numériques, je voudrais dire un mot d'une expérience fort intéressante dans laquelle le résultat se traduisait par un dessin. Il s'agit d'observations de la surface d'une planète, faites en un même instant, par des observateurs éloignés les uns des autres, dont chacun communiquait un dessin de son observation. Ces dessins ont été réunis et comparés par M. Jean Mascart, de l'Observatoire de Paris, à la très intéressante étude duquel je renvoie [1]. L'expérience suggérée par M. Nicolas Poutiata a été organisée par la Société Astronomique de France, sur la proposition de M. Camille Flammarion ; le programme en avait été arrêté avec soin et fut exécuté avec le concours de 36 observateurs, disposant de moyens d'action fort divers ; elle a duré 19 jours (du 2 au 20 janvier 1906), le mauvais temps ayant naturellement empêché certains jours le concours de certains observateurs ; mais le nombre des observations faites un même jour à la même heure dépasse souvent 10 et s'élève jusqu'à 17. Nous ne pouvons entrer ici dans la discussion détaillée à laquelle se livre M. Jean Mascart, ni dans la description des divers procédés photographiques employés pour *prendre la moyenne* des dessins divers (signalons toutefois l'ingé-

1. Observations simultanées de la surface de Jupiter, réunies par M. Jean Mascart (*Bulletin de la Société astronomique de France*, 1907).

nieuse idée de proportionner les temps d'action photographique, soit au diamètre, soit à la surface de l'objectif employé dans l'observation particulière utilisée). Contentons-nous de retenir l'impression qui résulte de la seule inspection des dessins juxtaposés : ces dessins exécutés le même jour à la même heure, sont très différents les uns des autres et fort peu de détails se retrouvent sur plusieurs d'entre eux. Cette simple constatation doit rendre très prudent dans l'interprétation de dessins de cette nature ; il est manifeste que des observateurs dont la bonne foi n'est pas douteuse, et qui savaient de plus qu'ils seraient soumis à un contrôle, ont cru voir des apparences ne correspondant à aucune réalité. Ces apparences sont-elles dues aux conditions atmosphériques ou à l'imagination de l'observateur, c'est ce que de nouvelles expériences pourraient peut-être déterminer. En tout cas, l'expérience mériterait d'être reprise, car elle n'est pas moins intéressante pour le psychologue qu'elle ne promet d'être profitable pour les progrès de l'astronomie planétaire ; dans ces nouvelles expériences, on tirera grand profit de maintes remarques judicieuses de M. Jean Mascart.

105. — Parmi les observations psychologiques qui ont été l'occasion des développements mathématiques se rattachant au calcul des probabilités, on doit mentionner particulièrement la méthode bizarrement dénommée méthode *des cas vrais et des cas faux* (*Richtig und falsch ; Right and wrong*). Cette méthode a donné lieu à de nombreux calculs, dans lesquels interviennent l'intégrale de Gauss et des formules qui s'y

rattachent ; ce n'est pas ici le lieu d'en faire un exposé critique, qui exigerait un appareil mathématique relativement considérable ; je me contenterai de signaler que le point délicat est de trouver une expression mathématique de l'erreur élémentaire satisfaisant également bien aux conditions de continuité et de discontinuité qui sont dans la nature de la question ; les tentatives faites dans ce sens ne me paraissent pas entièrement satisfaisantes au point de vue théorique et, d'autre part, conduisent rapidement, au point de vue pratique, à des complications analytiques qu'il serait préférable d'éviter, en des questions où sont intéressées beaucoup de personnes dont la culture mathématique est forcément limitée.

Je voudrais surtout indiquer comment l'emploi systématique de la méthode des majorités permet de poser la question sous une forme assez différente de la forme usuelle : au lieu de s'intéresser comme il semble qu'on l'ait fait surtout jusqu'ici, à l'étude de la sensation individuelle, qu'on cherchait à dégager de l'ensemble des expériences, c'est la précision de la « sensation collective » qu'il me paraîtrait intéressant de dégager de cet ensemble. Prenons un exemple précis.

L'un des buts que se sont proposés les expérimentateurs en psychologie, c'est de déterminer quel est l'accroissement minimum de poids que peut apprécier un observateur moyen ; étant donné un poids étalon de 1.000 grammes, on constate que c'est seulement lorsque le poids à juger est inférieur à 950 grammes ou supérieur à 1.050 que la différence est assez généralement

perçue [1] ; on dira donc que 50 grammes représentent la différence minimum perceptible par rapport à 1.000 grammes ; c'est de la détermination de cette différence minimum qu'on s'est beaucoup occupé depuis la loi de Weber. Mais on peut poser autrement la question ; au lieu de se proposer d'étudier la loi d'*une* expérience, on peut rechercher quelle précision pourraient donner un grand nombre d'expériences, et il est visible que cette précision serait plus grande, en même temps qu'il est évident qu'elle ne serait pas indéfinie. Supposons, par exemple, que l'on fasse comparer à un grand nombre d'observateurs deux poids, l'un de 1.000 grammes, l'autre de 1.010 grammes [2] ; si, sur 1.000 observateurs, il y en a 600 qui donnent la solution juste, il sera légitime de dire que cet ensemble de 1.000 observateurs a la perception collective de la différence de 10 grammes bien que, si l'on recommençait plusieurs fois l'expérience, il puisse arriver que chacun des 1.000 observateurs se trompe quelquefois, de sorte qu'aucun d'entre eux ne peut être regardé comme ayant la perception individuelle de cette différence. J'ai supposé ici que l'on exigeait une réponse affirmative ou négative, en excluant les réponses douteuses ; il serait aisé de poser le pro-

1. Je laisse ici entièrement de côté la discussion des erreurs systématiques qui s'introduisent dans les expériences de ce genre ; je suppose, par exemple, que l'on élimine l'erreur relative à l'ordre des comparaisons en faisant chaque comparaison deux fois ; je laisse aussi de côté la définition précise de l'écart *généralement perçu* ; si l'on admet que les écarts perçus suivent la loi de Gauss, on peut adopter, soit l'écart probable, soit l'écart médian, assez voisins l'un de l'autre, comme on sait.

2. Bien entendu, je suppose toujours que les erreurs systématiques soient évitées.

blème d'une manière analogue dans le cas où l'on admet ces réponses douteuses. Supposons, par exemple, que l'on ait une échelle de poids de comparaison, marqués en chiffres, et variant de gramme en gramme depuis 900 jusqu'à 1.100 grammes; on donne à l'observateur un poids P de valeur inconnue et on lui demande quels sont ceux des poids connûs qui lui paraissent égaux, lesquels supérieurs, lesquels inférieurs au poids à déterminer. Passant encore ici sur les détails de l'expérience, non que j'en méconnaisse la grande importance, mais je dois me borner, je suppose que l'on déduise de son résultat la valeur la plus probable du poids P [1]. Quelle confiance peut-on accorder à une telle détermination individuelle? Et quelle confiance peut-on attribuer à la moyenne de 1.000 déterminations analogues, faites par des observateurs différents? Si l'on acceptait d'une manière aveugle la loi de Gauss, ou toute autre loi mathématique précise, pour exprimer les erreurs individuelles, on serait forcément conduit à la conclusion que la précision croît comme la racine carrée du nombre des observations et peut, par suite, être rendue aussi grande que l'on veut. Cette méthode d'estimation en arriverait ainsi à pouvoir dépasser, si on l'employait avec soin, la précision des meilleures mesures faites au moyen d'instruments, ce qui est en apparence paradoxal [2]. Il serait intéressant

1. Les méthodes mathématiques pour atteindre ce but sont très simples si l'on suppose l'intervalle assez faible pour que l'on puisse y considérer le minimum perceptible comme constant; elles se compliquent si l'on tient compte de sa variation suivant la loi de Weber.

2. Le paradoxe s'atténue si l'on observe qu'il est pratiquement *impossible* de faire effectuer *la même* mesure à deux observateurs,

de déterminer par l'expérience la précision collective qui peut ainsi être atteinte par des mesures estimatives. C'est là un problème qui ne me paraît pas indigne de retenir l'attention des observateurs ; on trouverait d'ailleurs dans les publications déjà faites, et surtout sans doute dans les cahiers d'expériences, des éléments pour en commencer l'étude [1].

Pour donner une idée de la marche que l'on pourrait suivre, j'emprunte les chiffres suivants au *Manuel de psychologie expérimentale* de Titchener [2]. Un poids S de 1.071 grammes était comparé à divers poids C, au nombre de 7, dont les valeurs croissaient de 921 à 1.221 grammes par intervalles de 50 grammes. Le tableau suivant donne les résultats de 100 observations pour chaque poids C, soit 700 observations en tout ; on trouve dans les trois colonnes les nombres respectifs des observations où C fut jugé supérieur, égal, ou inférieur à S.

car la grandeur se modifie avec le temps (et aussi les instruments de mesure). Lorsque le nombre des observateurs devient très grand, ces variations en arrivent à dépasser le quotient de l'erreur individuelle moyenne par la racine carrée du nombre des observateurs ; à partir de ce moment, la précision ne peut plus être augmentée. La question est de savoir si cette précision limite peut être effectivement atteinte.

1. D'après des renseignements qui m'ont été fournis par M. Ch.-Ed. Guillaume, directeur-adjoint du Bureau international des Poids et Mesures, on est arrivé, en multipliant les lectures faites par des observateurs exercés, à atteindre dans les moyennes une précision *beaucoup plus* grande que la précision des observations individuelles : par exemple, au moyen de lectures faites à la loupe sur une échelle millimétrique, on a pu atteindre avec certitude *le centième de millimètre*, avec un nombre suffisant d'observations.

2. Titchener, *Experimental Psychology. Student's Manual-Quantitative*, p. 107 ; j'ai fait la somme des nombres correspondant aux deux séries d'expériences (Time Order I, S first et Time Order II, S second).

Valeur de C	C > S	C = S	C < S
921	1	8	91
971	6	11	83
1.021	10	37	53
1.071	30	37	33
1.121	55	34	11
1.171	76	20	4
1.221	90	9	1

Si l'on examine ce tableau au point de vue auquel nous nous plaçons, c'est-à-dire si l'on cherche à déduire de l'ensemble des observations la relation réelle supposée inconnue entre C et S, on aperçoit immédiatement que cette relation apparaît sans ambiguïté ; la lecture des trois dernières colonnes indique sans qu'il soit nécessaire de regarder la première si C est supérieur, égal, ou inférieur à S. Bien entendu, l'égalité rigoureuse ne peut pas être établie, mais simplement présumée d'après le résultat des expériences. Du fait que sur 100 expériences, C fut jugé 30 fois supérieur, 37 fois égal et 33 fois inférieur, on peut simplement conclure ceci : il n'y a aucune raison[1], d'après ces expériences, pour supposer $C > S$ plutôt que $C < S$; on doit donc, si l'on n'a pas d'autres éléments d'information, supposer $C = S$. De même, la majorité distingue à coup sûr un poids C de

1. En réalité, la petite différence entre 30 et 33 donnerait une faible raison de supposer $C < S$; c'est là un problème de probabilité des causes, qui peut être traité, suivant la méthode classique, à l'aide d'une hypothèse supplémentaire sur la probabilité *a priori*. Mais il n'est pas besoin de calcul, pour peu qu'on soit familier avec ces questions, pour être certain que la probabilité pour que l'on ait $C < S$ dépasse très peu 0,5, et est par suite très voisine de la probabilité pour que $C > S$.

1.021 grammes du poids S de 1.071 grammes, bien que chacun des observateurs puisse se tromper individuellement ; car, en laissant de côté les 37 réponses affirmant l'égalité ou le doute, il reste 53 réponses exactes contre 10 inexactes ; une telle répartition ne peut pas être l'effet du hasard ; si un jeu de pile ou face, sur 63 épreuves, donne 53 fois pile et 10 fois seulement face, la certitude est presque absolue que le jeu était truqué.

Mais on peut aller plus loin et se demander ce qui se passerait si C avait une valeur intermédiaire entre 1.021 grammes et 1.071 grammes ; vu la régularité de la variation des nombres inscrits dans les trois colonnes, on commet une erreur assez faible en procédant par interpolation linéaire, c'est-à-dire en supposant que les nombres varient en progression arithmétique, ce qui donne les résultats suivants [1] :

C	C > S	C = S	C < S
1.021	10	37	53
1.031	14	37	49
1.041	18	37	45
1.051	22	37	41
1.061	26	37	37
1.071	30	37	33

Bien entendu, si l'expérience était réellement faite, les nombres ne seraient pas exactement ceux que nous inscrivons à cause des erreurs for-

1. En tenant compte des différences secondes, on obtient des nombres assez voisins qui conduisent aux mêmes conclusions finales j'ai préféré présenter le calcul sous sa forme la plus intuitive.

tuites; mais la théorie des probabilités nous indique quelles sont les erreurs fortuites qui peuvent être regardées comme probables et celles qui sont hautement improbables. Par exemple, sur 63 jugements (ce qui est le cas ici, puisqu'il y en a 37 neutres), on doit regarder, s'ils sont l'effet du hasard, comme peu probable qu'il y en ait plus de 40 dans un même sens, 23 seulement étant de sens opposé [1]. L'ensemble des observations faites dans le cas où $C = 1.051$ permet donc de conclure avec beaucoup de vraisemblance que l'on a $C < S$, c'est-à-dire dans le sens de la vérité. De l'ensemble des nombres de notre premier tableau, nous tirons donc la conclusion que l'ensemble des 100 observateurs considérés, lorsque l'on considère leur jugement collectif, arrive à distinguer une différence de poids de 20 grammes. Or, par des calculs assez longs et très corrects Titchener (*loc. cit.*, p. 109 et p. 113) donne pour les valeurs de DL[2] déduites de ce même tableau les divers nombres suivants [3]: 43, 36, 30, 53, 38, 34, dont la moyenne 39 est à peu près le double du nombre 20 que nous trouvons. Cette différence des résultats déduits des mêmes expériences montre bien que les problèmes sont distincts :

1. La probabilité est $1 - \Theta\left(\frac{9}{\sqrt{32}}\right) = 1 - \Theta(1,6) = 0,024$ (Θ désignant l'intégrale de Gauss).

2. C'est-à-dire du minimum d'accroissement moyen perceptible pour l'observateur individuel.

3. Les différences entre ces nombres sont dues, d'une part à la distinction entre les deux « time order », d'autre part à la distinction entre les accroissements positifs et négatifs, et, enfin, à des différences entre les méthodes suivies ; ces différences entre les résultats montrent combien il est illusoire d'appliquer des méthodes de calcul trop précises à des nombres grossièrement approchés.

Titchener détermine la sensibilité individuelle tandis que nous déterminons la sensibilité collective de l'ensemble des observateurs. Il serait d'ailleurs possible de donner plus de précision à notre détermination, mais ce serait aux dépens de la rigueur, vu la petitesse des nombres employés [1]. Il vaut mieux s'en tenir à cette constatation nette, que la différence de 20 grammes est perceptible presque à coup sûr par la méthode des majorités appliquée à l'ensemble des 100 observations ; cette méthode donne donc, comme on pouvait s'y attendre, une précision et une sécurité bien plus grande que l'observation individuelle.

Je me contenterai de cet exemple, espérant qu'il suffira pour montrer quel intérêt il pourrait y avoir à appliquer la méthode à des expériences combinées spécialement en vue de cette application. Il n'est pas douteux qu'on serait ainsi conduit à des résultats importants à divers titres.

106. — Pour terminer, je voudrais discuter quelques applications de la méthode des majorités à des questions qualitatives, car c'est peut-être là surtout que cette méthode fournit des résultats qui ne pourraient être obtenus par aucun autre moyen.

1. Si l'on suivait une méthode parallèle à celle qui est généralement employée dans ces questions quand on use de l'intégrale de Gauss, on rechercherait quelle valeur doit être donnée à C pour que la probabilité de l'écart observé soit inférieure à $\frac{1}{2}$; on déduirait ainsi des nombres donnés que la valeur de DL, pour l'ensemble des 100 observateurs, considérés collectivement, est d'environ 8 grammes ; mais le nombre des observations est trop faible pour que ce résultat puisse être regardé comme précis.

La première idée de ce genre d'application m'a été suggérée par une enquête que j'ai faite sur des textes qui avaient servi à des expériences graphologiques de M. Binet [1]. Il s'agissait de distinguer, sur 12 textes très courts, reproduits en typographie, lesquels des 12 scripteurs étaient des hommes supérieurs et lesquels des hommes médiocres [2]. Les réponses individuelles furent intéressantes, mais le classement déduit du dépouillement du scrutin le fut plus encore; les six hommes supérieurs furent déclarés tels par la majorité absolue; un seul homme médiocre eut aussi la majorité absolue et c'est, si l'on peut dire, le moins médiocre des hommes « médiocres » (ou « simplement intelligents ») choisis par M. Binet, qui en dit: « C'est un publiciste scientifique qui ne manque pas de mérite et qui, s'il se reconnaît, ne m'en voudra pas de lui avoir préféré Claude Bernard. » Somme toute, l'opinion de la majorité était conforme à la vérité pour tous les hommes supérieurs et l'erreur unique qu'elle commettait en classant comme supérieur un des hommes médiocres pouvait se justifier par les renseignements mêmes donnés sur lui. Je n'insisterai pas plus longtemps sur cette expérience, mais j'ai tenu à la rappeler, car ce fut à son occasion que je fus amené à correspondre avec M. Binet au sujet de cette méthode des majorités (c'est lui qui employa le premier cette expression), et qu'il voulut bien m'engager à écrire cette petite étude pour son *Année*.

1. Voir la *Revue du Mois* d'août et de septembre 1906, t. II, p. 244 et 366.

2. La classification était due à M. Binet; voir son livre : *Les Révélations de l'écriture soumises à un contrôle scientifique* (F. Alcan).

Je discuterai en détail une expérience intéressante faite par Mme Rousson sur la lecture d'une physionomie d'enfant [1]. Rappelons-en brièvement les conditions, en laissant de côté les détails. On soumit séparément à 20 observateurs 40 photographies d'enfants (chacune double ; face et profil), dont 23 étaient arriérés et 17 normaux ; chaque observateur déclarait l'enfant arriéré, ou normal, ou refusait de répondre (il y eut seulement 18 refus de réponse sur 800 observations [2]). Les pourcentages de réponses justes varièrent suivant les observateurs, entre 67 p. 100 et 92 p. 100 ; leur moyenne fut de 78 p. 100. On peut conclure de là qu'il y avait vraiment dans les photographies [3], ou du moins dans certaines d'entre elles, des indices d'intelligence. Mais cette conclusion générale peut être de beaucoup précisée par l'application de la méthode des majorités. Je voudrais montrer en détail sur cet exemple quels principes simples on doit utiliser dans cette application, en tâchant de faire appel au minimum de connaissances mathématiques.

1. *Bulletin de la Société libre pour l'étude psychologique de l'enfant*, juin 1906 (6e année, p. 76).

2. Ces refus sont assez peu nombreux pour pouvoir être négligés. Mme Rousson n'en tient pas compte dans le pourcentage des réponses justes, c'est-à-dire que si un observateur refuse de répondre 2 fois et se trompe 3 fois, sur 40 enfants, elle compte 35 réponses justes sur 38. Ce procédé augmente légèrement les pourcentages de Mme Rousson ; nous les avons reproduits tels qu'elle les donne.

3. Je dis les *photographies*, et non pas seulement les *physionomies*, car un des observateurs a fait remarquer qu'il eût mieux valu « écarter toutes les différences extérieures qui peuvent influencer le jugement ». Cette remarque est sans importance pour nos déductions ; elle doit intervenir dans les conclusions spéciales à la physionomie que l'on pourrait chercher à en tirer.

Le nombre des votants était 20; chacun d'eux devait voter : *normal* (N) ou *arriéré* (A), de sorte que l'ensemble des votes relatifs à une même photographie était représenté par une suite telle que la suivante [1] (photographie n° 7) :

NNAAANANAAANAAAAAAAA

Combien peut-il exister de telles suites de 20 lettres N ou A ? La première lettre peut être N ou A, ce qui fait deux possibilités ; pour chacune de ces 2 possibilités, la seconde lettre peut être N ou A, ce qui fait 2×2 possibilités (NN, NA, AN, AA); pour chacune de ces 2×2 possibilités, la troisième lettre peut être N ou A, ce qui fait $2 \times 2 \times 2$ possibilités (NNN, NNA, NAN, NAA, ANN, ANA, AAN, AAA), etc. En répétant le même raisonnement jusqu'à la vingtième lettre, on trouve un nombre de possibilités égal au produit de 20 facteurs égaux à 2, c'est-à-dire à 2^{20}, ou à 1.048.576. Tel est le nombre des diverses possibilités du scrutin considéré. Si chaque votant, au lieu de délibérer son vote, le jouait à pile ou face, chacune de ces possibilités serait également probable, et il y aurait par suite exactement 1 chance sur 1.048.576 pour que l'une quelconque d'entre elles se produisît (par exemple celle que nous avons écrite plus haut).

1. Je laisse de côté les quelques votes *nuls* (au nombre de 18 sur 800) ; pour donner plus de certitude à mes conclusions, je les compterai avec la *minorité*; c'est le procédé le plus désavantageux et avec toute autre convention, les conclusions seraient vraies *a fortiori*. Vu leur petit nombre, la chose est d'ailleurs de peu d'importance.

Il faut maintenant déterminer combien, parmi les diverses possibilités du scrutin, il y en a qui donnent pour N un certain nombre de voix ; il est bien clair, par exemple, que N peut avoir une voix contre 19 à A de 20 manières différentes, car l'unique voix N peut être celle du premier votant, ou celle du second,..... ou celle du vingtième. La solution complète de cette question est fournie par l'analyse combinatoire et le calcul numérique se fait simplement par l'algorithme dit « triangle arithmétique de Pascal » ; je me contenterai d'en indiquer le résultat.

On peut obtenir	D'un nombre de manières égal à	Somme des nombres de la colonne précédente.
20 N et 0 A	1	1
19 N et 1 A	20	21
18 N et 2 A	190	211
17 N et 3 A	1.140	1.351
16 N et 4 A	4.845	6.196
15 N et 5 A	15.504	21.700
14 N et 6 A	38.760	60.460
13 N et 7 A	77.520	137.980
12 N et 8 A	125.970	263.950
11 N et 9 A	167.960	431.910
10 N et 10 A	184.756	616.666
9 N et 11 A	167.960	....
8 N et 12 A	125.970	....

Il a été inutile d'écrire la fin du tableau, qui reproduit évidemment le commencement en ordre inverse. Dans la troisième colonne, nous avons inscrit la somme des nombres de la seconde (21.700, par exemple, est la somme 1 + 20 + 190 + 1.140 + 4.845 + 15.504) ; on verra tout à l'heure quelle en est l'utilité.

Le tableau précédent est l'élément essentiel de

la discussion de la méthode des majorités ; lorsque le nombre des votants est petit, on l'obtient sans peine à l'aide du triangle arithmétique de Pascal ; lorsqu'il est grand, les calculs seraient inextricables et l'on utilise des formules approchées que l'on établit par le calcul intégral et qui conduisent à considérer la célèbre intégrale de Gauss ; telle est la signification de cette intégrale, qui n'a rien de mystérieux ; mais dans l'exemple actuel, il m'a paru préférable d'en éviter l'emploi.

Le nombre total des possibilités est 1.048.576 ; si les votes étaient joués à pile ou face, ce qui rendrait toutes ces possibilités également probables, la probabilité pour que se produise l'une des 38.760 combinaisons qui donne 14 N et 6 A serait égale au quotient de 38.760 par 1.048.576, c'est-à-dire, en chiffres ronds, à 4 p. 100. On peut ainsi déduire du tableau précédent les résultats probables que donneraient des votes émis au hasard ; les voici, en arrondissant les chiffres :

On obtiendrait	20 N et 0 A	1	fois sur	1.000.000
—	19 N et 1 A	1	—	50.000
—	18 N et 2 A	1	—	2.500
—	17 N et 3 A	1	—	1.000
—	16 N et 4 A	1	—	200
—	15 N et 5 A	1.5	—	100
—	14 N et 6 A	4	—	—
—	13 N et 7 A	8	—	—
—	12 N et 8 A	12	—	—
—	11 N et 9 A	17	—	—
—	10 N et 10 A	18	—	—

Il en résulte que, sur 40 scrutins, on devrait s'attendre, en gros, aux résultats suivants :

7 fois chacune des combinaisons	11 N et 9 A, 10 N et 10 A 9 N et 11 A
5 — —	12 N et 8 A, 8 N et 12 A
3 — —	13 N et 7 A, 7 N et 13 A
3 fois en tout les deux combinaisons	14 N et 6 A, 6 N et 14 A

l'arrivée des autres combinaisons étant, non pas impossible, mais assez peu probable pour 15 N ou 16 N, très peu probable pour 17 N et 18 N, tout à fait improbable pour 19 N ou 20 N. En particulier, cette dernière combinaison (l'unanimité) ne devrait se produire que 1 fois sur 1.000.000, or, en réalité, elle s'est produite 5 fois et chaque fois dans le sens de la vérité [1]. Attribuer cet événement au hasard est aussi improbable qu'admettre le fait suivant : dans une ville comme Londres, renfermant environ 1.000.000 d'hommes adultes, 40 individus investis du pouvoir suprême sont forcés de démissionner pour remettre le pouvoir à 5 citoyens tirés au sort parmi le million, et le tirage au sort désigne précisément 5 des démissionnaires. Le peuple sera sûr qu'il y a fraude ; et il aura raison. De même, nous devons être certains que ce n'est pas le hasard seul qui a produit le résultat du vote ; il y a donc une cause, et il est assez naturel de penser que c'est bien la cause recherchée, c'est-à-dire l'inspection des photographies (ce pourrait être aussi une entente entre les votants, ou une suggestion ; ces hypothèses paraissent devoir être écartées d'après ce qui est rapporté des conditions de l'expérience).

1. C'est parce qu'elle s'est produite dans le sens de la vérité que nous disons 1 fois sur 1.000.000 et non 2 fois sur 1.000.000, qui serait la probabilité pour que se produise l'unanimité dans un sens non indiqué d'avance (soit 20 N, soit 20 A).

Mais entrons dans le détail des résultats ; nous les grouperons en exacts, douteux et inexacts.

I. Résultats exacts[1].

5 à l'unanimité de 20 voix.
5 par 19 voix contre 1 (ou contre 1 abstention).
6 par 18 voix contre 2 (ou abstentions).
2 par 17 voix contre 3 (—).
4 par 16 voix contre 4 (—).
3 par 15 voix contre 5 (—).
2 par 14 voix contre 6 (—).

27 résultats exacts au total.

II. Résultats douteux.

2 à la majorité de 13 voix.
2 — 12 —
1 — 11 —
4 à l'égalité des voix (10 contre 10).
1 à la minorité de 9 voix (contre 10 et 1 abstention).
1 — 5 voix[2] (contre 12 et 3 abstentions).

11 douteux.

1. Ces résultats exacts ou douteux se répartissent à peu près également entre les normaux et les arriérés ; voici cette répartition :

I. Exacts.		II. Douteux.	
20 voix	4 A, 1 N	13 voix	1 A, 1 N
19 —	2 A, 3 N	12 —	1 A, 1 N
18 —	5 A, 1 N	11 —	0 A, 1 N
17 —	0 A, 2 N	10 —	3 A, 1 N
16 —	1 A, 3 N	9 —	0 A, 1 N
15 —	3 A, 0 N	5/8 —	1 A, 0 N
14 —	0 A, 2 N		

Les deux résultats inexacts concernent deux arriérés déclarés normaux par le vote.

2. Ce résultat aurait pu aussi être rangé parmi les inexacts ; il est douteux si l'on ajoute les abstentions à la minorité.

III. Résultats inexacts.

1 où la solution exacte a seulement 5 voix contre 15.
1 — — 3 — 17.

2 inexacts.

Les résultats classés comme douteux sont ceux pour lesquels *il n'est pas improbable que l'exactitude du résultat soit due au hasard.* Leur nombre est 11; si l'on se reporte à la troisième colonne du tableau de la page 277 on voit que sur 1.048.576 épreuves régies par le hasard il y en a 60.460 pour lesquelles la majorité atteint ou dépasse 14 et 137.980 pour lesquelles cette majorité atteint ou dépasse 13; donc, sur 11 épreuves régies par le hasard, le nombre probable de celles pour lesquelles la majorité dans un sens donné d'avance serait égale ou supérieure à 14 est inférieur à 1, tandis que le nombre de celles pour lesquelles la majorité atteint 13 est supérieure à 1. C'est pour cela que nous avons classé les 2 scrutins où la majorité atteint 14 parmi les résultats exacts et les 2 scrutins où la majorité atteint 13 parmi les douteux. Il y a évidemment un certain arbitraire dans cette démarcation absolue; il serait puéril de dissimuler cet arbitraire; mais il serait excessif d'en tirer argument contre les conclusions dans ce qu'elles ont de général; on doit simplement ne pas oublier que ces conclusions comportent quelque jeu, c'est-à-dire qu'au lieu de 27 résultats exacts sur 40, on pourrait en compter 25 à 29, peut-être même 24 à 30; mais cette légère incertitude dans la précision numérique n'entache pas le principe même des déductions: il est dans la nature même de tout calcul

basé sur des expériences de ne fournir que des résultats approchés. Ces réserves faites, nous adopterons comme exacts les nombres des pages 280 et 281.

On voit que, pour 27 photographies sur 40, c'est-à-dire pour *les deux tiers* environ, on est conduit à admettre que le résultat du scrutin a été déterminé par l'aspect des photographies ; il est tout à fait vraisemblable que ces 27 photographies, soumises à d'autres observateurs de même nature (instituteurs et institutrices), auraient conduit aux mêmes résultats ; il serait intéressant de tenter l'expérience.

Pour 13 photographies, c'est-à-dire pour *un tiers*, l'ensemble des réponses ne donne pas une exactitude supérieure à ce qu'aurait fourni le hasard. On doit, semble-t-il, en conclure que, pour ces photographies, les caractères sur lesquels se basaient en général les observateurs pour juger étaient absents ou partiellement contradictoires. En résumé, il semblerait que, pour un tiers des photographies, l'expérience doive être interprétée comme conduisant à la conclusion que les caractères de ces photographies *ne permettent pas* de déterminer à quelle catégorie appartient l'enfant. Telle est la conséquence logique de cette expérience; cette conséquence pourrait être partiellement modifiée par une expérience plus étendue portant sur les mêmes photographies ; je suis convaincu toutefois qu'elle ne serait pas entièrement infirmée.

Enfin, disons quelques mots des deux scrutins dont le résultat est inexact, auxquels on doit peut-être joindre l'un de ceux qui ont été classés comme douteux. Il s'agit de trois enfants arriérés

qui ont été déclarés normaux par 17 voix contre 3, 15 contre 5 et 12 contre 5. Il y a d'assez sérieuses présomptions pour que deux tout au moins des photographies (sinon les trois) possèdent les caractères des enfants normaux, c'est-à-dire pour que l'erreur de la majorité corresponde à une vérité relative, bien que contredisant la vérité absolue (c'est-à-dire l'appréciation des personnes qui connaissent très bien les enfants). On pourrait se demander si ce n'est pas parce que ces photographies sont plus difficiles à juger qu'elles ont ainsi conduit à des conclusions inexactes, et si des juges plus habiles n'arriveraient pas, sur elles, à la vérité. Il ne le semble pas, car ceux des observateurs qui ont donné sur ces photographies une appréciation exacte ne paraissent pas se distinguer par une habileté particulière ; les pourcentages respectifs de réponses exactes sont, pour les observateurs ayant bien répondu dans les trois scrutins à résultat inexact :

82 p. 100, 75 p. 100, 67 p. 100, 78 p. 100, 67 p. 100
82 p. 100, 67 p. 100, 92 p. 100
77 p. 100, 76 p. 100, 67 p. 100, 82 p. 100, 75 p. 100

Ces pourcentages oscillent entre le minimum général et le maximum général des 20 observateurs (67 p. 100 et 92 p. 100) ; leur moyenne, 76 p. 100, est légèrement inférieure à la moyenne générale, 78 p. 100, des 20 observateurs. Ce ne sont donc pas les observateurs les plus habiles qui ont le mieux répondu dans ces cas-là.

Comme conclusion générale de cette application de la méthode des majorités à l'expérience de Mme Rousson, nous arrivons à la conclusion que les photographies renfermaient les éléments

d'une solution exacte dans environ 65 p. 100 des cas[1], devaient laisser dans le doute dans environ 30 p. 100 des cas, et devaient conduire à une solution inexacte (arriéré pris pour normal) dans environ 5 p. 100 des cas.

107. — J'ai appliqué une critique analogue à la précédente à une expérience fort intéressante de M. Binet sur des photographies de mains[2]. Il s'agissait de déterminer, sur une photographie de main d'enfant (ou plutôt deux photographies : paume et dos) 1° quel était le sexe de l'enfant ; 2° si l'enfant était intelligent ou bête. La réponse à cette seconde question posée sous cette forme par M. Binet devait être donnée sans ambiguïté. L'expérience a porté sur 20 photographies, qui ont été soumises à 20 observateurs ; j'ai posé moi-même à 4 observateurs la question de M. Binet relative à l'intelligence, et j'ai joint dans ce qui suit ces 4 réponses aux 20 obtenues par M. Binet[3] ; voici les résultats :

1° Question relative au sexe :

13 réponses exactes :

2 par 19 voix sur 20 votants, 1 par 18 voix, 6 par 17 voix, 1 par 16 voix, 1 par 15 voix, 1 par 14 voix.

1. Ceci ne contredit pas le fait que le pourcentage général des réponses exactes a été de 78 p. 100, car, dans les cas douteux, il y a environ une chance sur deux pour que la réponse soit exacte ; si tous les cas étaient douteux, le pourcentage général des réponses exactes serait de 50 p. 100.

2. Cette expérience est inédite : je remercie M. Binet de m'en avoir communiqué le dossier.

3. Cette adjonction diminue d'une unité le nombre des *douteux* ; le sens général du résultat reste le même. Dans le dossier que m'a communiqué M. Binet, j'ai utilisé, pour chaque observateur, les réponses de la première épreuve, sans tenir compte des modifications de jugement à la seconde épreuve, dont je dirai un mot tout à l'heure.

5 réponses douteuses :

1 par 13 voix, 2 par 12, 1 par 11, 1 par 7.

2 réponses inexactes :

5 voix pour la vérité contre 15.

Les conclusions sont à peu près semblables à celles que nous avons tirées de l'expérience de Mme Rousson, sauf, toutefois, que les majorités dans les réponses exactes sont moins considérables ; les photographies des mains dévoilent un peu moins bien le sexe que les photographies des physionomies ne dévoilent l'intelligence. A signaler que les deux réponses inexactes se rapportent à deux garçons pris pour des filles : ces deux garçons sont intelligents : l'un d'eux a été jugé tel par 18 voix sur 24 et l'autre a eu dans cette seconde enquête 12 voix sur 24.

2° Question relative à l'intelligence :

6 réponses exactes :

1 par 19 voix sur 24, 2 par 18, 1 par 17 et 1 par 16;

Ces réponses exactes concernent 2 enfants jugés intelligents et 4 jugés bêtes.

12 réponses douteuses :

La solution exacte ayant obtenu, sur 24 votants, 3 fois 14 voix, 4 fois 12 voix, 2 fois 11 voix, 2 fois 10 voix, 1 fois 9 voix.

2 réponses inexactes :

La solution exacte n'ayant obtenu que 7 voix sur 24.

Ces deux réponses inexactes se réfèrent respectivement aux deux catégories.

Ces résultats sont, comme il était naturel de

s'y attendre, très inférieurs comme exactitude à ceux qu'a fournis l'étude des physionomies; la majorité n'est jamais aussi nette et, dans les deux tiers des cas, les écarts sont de l'ordre de ceux qu'on pourrait attendre du hasard. Il semble donc bien qu'il y ait environ 60 p. 100 de cas dans lesquels on ne peut rien tirer des photographies de mains. Ce qui corrobore cette impression, c'est l'étude d'une seconde expérience, dans laquelle un certain nombre des mêmes observateurs ont varié dans leur jugement sur les mêmes photographies (dans cette seconde épreuve, ils connaissaient le sexe des enfants); ces variations de jugement sont au total de 62, dont 46 sur les 12 cas que nous avons classés comme douteux, 15 sur les 6 cas où la réponse fut exacte et 1 seulement sur les deux cas où la réponse fut inexacte. Cela donne environ 4 variations pour chacun des cas douteux, 2,5 seulement pour les réponses exactes et 0,5 pour les réponses inexactes. Il y a donc des cas où le jugement est très incertain, et d'autres où il est plus sûr, et où l'examen décèle « quelque chose d'objectif ». Ce « quelque chose d'objectif » s'est trouvé coïncider avec l'intelligence dans 6 cas sur 8; ces chiffres sont trop faibles pour que le calcul des probabilités permette de rien dire à leur sujet; il serait illusoire de faire des pourcentages; en remplaçant 6 sur 8 par 75 p. 100, on commettrait une grave erreur; car il peut fort bien arriver que le hasard pur (c'est-à-dire une probabilité de 50 p. 100) donne 6 succès sur 8 expériences, tandis que 75 succès sur 100 expériences décèlent, avec une certitude pratiquement absolue, une cause.

C'est donc par une étude concrète que l'on peut

chercher à déterminer la nature de ce « quelque chose d'objectif » que décèle l'expérience. L'examen des photographies a suggéré à M. Binet l'hypothèse suivante : lorsque les doigts sont courts et laids, on juge l'enfant bête; lorsqu'ils sont longs et effilés, on le juge intelligent. Pour vérifier cette hypothèse, M. Binet a demandé à deux nouveaux expérimentateurs de faire une classification en basant leur jugement uniquement sur le critérium précédent. Cette expérience figure dans le dossier qu'il m'a remis; pour ne pas me laisser influencer par elle, je ne l'ai dépouillée qu'après avoir rédigé tout ce qui précède : elle en apporte une confirmation remarquable, de sorte que l'hypothèse de M. Binet me paraît entièrement confirmée par mon étude arithmétique. En effet, dans les 6 cas où la réponse de la majorité a été exacte, et dans les 2 cas où elle a été inexacte, *toutes* les réponses de ces deux derniers observateurs concordent exactement avec celles de la majorité [1]. Cette concordance parfaite, portant sur 16 opinions (2 observateurs jugeant chacun 8 photographies), est en quelque sorte la vérification expérimentale de la méthode des majorités, et cet exemple montre bien comment cette méthode devra être appliquée dans les cas, comme celui-ci, où l'on ne sait pas très exactement *a priori* sur quoi porte la recherche. La méthode des majorités nous apprend qu'il y a « quelque chose d'objectif »; l'observation directe des faits conduit en-

1. Une seule différence insignifiante à signaler; pour l'un des enfants jugé bête par la majorité (réponse exacte), l'un des deux observateurs a inscrit « moyen », l'autre ayant d'ailleurs inscrit « bête ».

suite à formuler des hypothèses, qui doivent être vérifiées par de nouvelles expériences. Le calcul n'a joué là, comme toujours dans les sciences expérimentales, qu'un rôle d'intermédiaire entre deux expériences, mais ce rôle de guide de l'expérience, pour être modeste, est fort souvent des plus utiles.

108. — Il serait possible de multiplier les exemples ; je crois qu'il serait intéressant et utile d'étudier systématiquement, par les méthodes qui viennent d'être esquissées, la masse considérable de documents rassemblés par les psychologues, en France particulièrement par Alfred Binet, ses collaborateurs et ses élèves (notamment dans l'*Année psychologique* et dans le *Bulletin pour l'étude psychologique de l'enfant*). Mais cette longue étude n'est pas indispensable pour que nous soyons dès à présent assurés de la valeur scientifique que peut avoir dans ces domaines la théorie des probabilités, appliquée avec rigueur et avec prudence.



CHAPITRE X

LA PORTÉE PHILOSOPHIQUE DES LOIS DU HASARD

109. L'étude des lois du hasard renouvelle le problème de la valeur de la connaissance. — 110. Le déterminisme du savant. — 111. Le déterminisme à l'échelle humaine n'entraîne pas le déterminisme *absolu* à l'échelle moléculaire. — 112. Le déterminisme à l'échelle moléculaire n'entraîne pas le déterminisme *absolu* à l'échelle humaine. — 113. Le miracle de la conscience épiphénomène est plus invraisemblable encore que le miracle des singes dactylographes. — 114. L'accord du monde avec la raison et l'évolution humaine. — 115. L'explication statistique de la loi de Newton serait regardée comme un progrès. — 116. Le second principe de la thermodynamique et la dégradation de l'énergie. — 117. Le « démon de Maxwell » et l'activité humaine. — 118. Une conception possible de l'évolution de l'univers.

109. — Ce n'est pas sans hésitation que je me suis décidé à écrire ce dernier chapitre ; à la réflexion, je me suis cependant convaincu que je ne pouvais me refuser à donner mon avis sur des questions qui doivent intéresser la plupart des lecteurs de ce livre. Il n'est pas besoin d'observer que, dès qu'il est question de la valeur philosophique d'une connaissance particulière, la question de la valeur de *la* connaissance se trouve nécessairement posée ; c'est assez dire qu'il ne peut être question d'apporter ici des conclusions définitives sur un problème qui n'a cessé de préoccuper l'homme depuis les origines de la pensée philosophique. Je voudrais essayer de montrer

BOREL. — 19

dans quelle mesure ce problème est renouvelé par l'étude des lois du hasard et par la constatation de la place qu'elles prennent dans la vie pratique et dans la connaissance scientifique.

110. — Henri Poincaré a fait observer que le savant ne peut pas, en tant que savant, ne pas être déterministe, de même que l'homme ne peut pas, dans la vie pratique, agir comme s'il ne croyait pas à l'existence du monde sensible. Le but de la science étant de prévoir, du moment où la prévision cesse d'être regardée comme possible, on est en dehors des frontières de la science, le savant cesse d'agir ou de penser en savant. Cette conception de la science est évidemment nécessaire ; toute la question est de savoir si cette nécessité est *absolue*, au sens des vérités mathématiques, ou si elle comporte une fraction de contingence, si faible qu'elle puisse être. A ce point de vue, les explications basées sur la théorie des probabilités, en particulier les explications « statistiques » des phénomènes physiques, sont doublement intéressantes à étudier : d'une part, elles permettent de comprendre que la nécessité d'un phénomène global n'est pas incompatible avec la « liberté » des phénomènes partiels ; d'autre part, elles fournissent des exemples de cas dans lesquels le déterminisme supposé absolu des phénomènes partiels ne permet pas de prévoir avec une rigueur *absolue* le phénomène global.

111. — Insistons un peu sur ces deux points, et tout d'abord sur le fait que la nécessité observée dans les « effets » n'entraîne pas nécessairement le déterminisme absolu des causes.

Nous avons trouvé, dans des sciences diverses, une même tendance : remplacer l'étude de phénomènes particuliers par l'étude globale, statistique, d'un ensemble de phénomènes très nombreux. Cette étude statistique conduit d'ailleurs à des résultats importants et précis, qui méritent véritablement le nom de lois scientifiques. Ce fait vaut qu'on y réfléchisse.

A la suite des grands savants de la fin du XVIIIe siècle, des Lagrange, des Laplace, des Fourier, le XIXe siècle s'est efforcé d'expliquer le plus grand nombre possible de phénomènes en les ramenant aux lois rigoureuses de la mécanique ; de même que le cours des astres est immuablement fixé et que tous les détails en peuvent être prévus, il devait suffire de perfectionner suffisamment les méthodes de calcul et d'analyser assez en le détail les faits physiques pour arriver à y constater le même accord entre la théorie et l'observation.

Il serait injuste de dire que cette conception a fait faillite, car c'est seulement en s'appuyant sur elle qu'on a pu la dépasser, et elle devra être conservée comme un des éléments indispensables à notre connaissance de l'univers ; mais, si elle reste nécessaire, elle n'est plus suffisante. Il est inutile de redire les raisons pour lesquelles l'analyse détaillée du déterminisme des phénomènes dépasse les forces humaines ; mais si les innombrables équations différentielles qui définiraient les mouvements des molécules ne peuvent être considérées individuellement, il est possible d'en étudier quelques propriétés générales, d'où l'on déduit précisément les lois statistiques. Le déterminisme rigoureux qui se déduirait de la résolution, théo-

riquement possible, de l'ensemble des équations ne joue dans l'établissement de ces lois statistiques qu'un rôle accessoire; ces lois resteraient les mêmes, si le détail des phénomènes était modifié, pourvu que les moyennes subsistent: parmi le milliard de milliards de molécules d'un gaz, quelques milliards choquent en un temps très court une portion de la paroi; il importe peu de savoir *quelles* sont ces molécules privilégiées, et, au point de vue humain, cette question n'a même aucun sens; la seule chose qui importe, c'est de connaître leur nombre et leur vitesse moyenne.

On pourrait aller plus loin, et imaginer que, parmi les milliards de milliards de molécules qui composent une petite masse gazeuse, quelques-unes soient soumises à des lois absolument différentes des lois que nous connaissons, ou même ne soient soumises à aucune loi et se comportent comme si elles étaient douées de libre arbitre au sens le plus naïf de ce terme; le déterminisme des phénomènes accessibles à nos observations n'en serait pas affecté; les lois physiques n'auraient à subir aucune modification[1]. De même,

1. Contre l'hypothèse que les molécules sont « libres », je dois citer un article fort suggestif de M. J.-H. Rosny aîné, *La contingence et la détermination*, paru pendant que je corrigeais les dernières épreuves de ce livre (*Revue du Mois*, 10 janvier 1914. t. XVII, p. 5). Je dois citer aussi une conférence très intéressante de M. Langevin, faite à la Société française de physique à la fin de novembre 1913 et encore inédite au moment où ce livre paraît. Dans cette conférence, M. Langevin a notamment insisté sur le fait que l'étude approfondie de certains phénomènes à notre échelle permet de fixer avec certitude certaines lois des phénomènes moléculaires : on ne peut se contenter de l'affirmation vague qu'ils suivent les lois de la théorie des probabilités. Mais, d'une part, cette détermination ne peut jamais être absolue et, d'autre part, on peut toujours concevoir que certaines molécules y échappent entièrement.

l'hypothèse que les individus sont doués du libre arbitre n'influe pas sur les lois de la statistique : le nombre des voyageurs qui quitteront Paris le 3 août par la gare Saint-Lazare peut être approximativement prévu ; la plupart d'entre eux cependant, considérés individuellement, pourraient aussi bien partir la veille ou le lendemain.

On pourrait multiplier les exemples ; mais je crois avoir suffisamment mis en évidence ce point essentiel : il n'y a pas incompatibilité entre le rôle de ce que nous appelons le hasard et l'établissement de lois scientifiques ; au contraire, en bien des cas, sinon dans tous les cas, les lois scientifiques sont la résultante simple d'un nombre extrêmement grand de phénomènes complexes, impossibles à étudier en détail. Cette conception rend-elle légitime l'application de méthodes proprement scientifiques à l'étude des phénomènes économiques et sociaux ? Ceux que cette question intéresse pourraient peut-être tirer quelque profit de la forme nouvelle sous laquelle on est conduit à envisager le déterminisme des phénomènes physiques et biologiques et qu'on peut appeler le déterminisme statistique.

Pour en revenir aux sciences physiques, le fait que le déterminisme global n'exclut pas l'hypothèse de l'indétermination des phénomènes à l'échelle moléculaire mérite d'être approfondi. Les fondateurs de la théorie cinétique ont pris, en effet, pour point de départ le déterminisme des phénomènes moléculaires ; n'y a-t-il pas quelque contradiction à mettre en doute ce déterminisme, en vertu des conséquences mêmes déduites de ce point de départ ? Nous reviendrons sur ce point tout à l'heure.

112. — Auparavant, disons quelques mots du fait en quelque sorte inverse du précédent, le caractère partiellement contingent des conséquences que l'on peut déduire de l'hypothèse du déterminisme *absolu* des phénomènes moléculaires. Il est bien clair que si les mouvements des molécules d'un gaz sont rigoureusement déterminés, toutes les propriétés de ce gaz pourraient être déduites avec une certitude absolue de l'étude des équations qui définissent les mouvements des molécules. Mais, même en laissant pour un instant de côté la difficulté, pourtant très réelle, due au nombre prodigieux d'équations qu'il faudrait écrire et intégrer, nous avons constaté qu'il n'est pas possible d'étudier avec précision le mouvement d'une molécule sans faire entrer dans les équations l'univers tout entier. Ce n'est donc plus seulement un nombre d'équations tellement grand que l'esprit humain est incapable de se les représenter toutes, c'est un nombre réellement infini d'équations qu'il faudrait supposer écrites et intégrées [1]. Le déterminisme du phénomène global peut donc bien être conçu d'une manière abstraite, mais le détail ne peut pas en être prévu ; tout ce que nous pouvons faire, c'est, par les méthodes de la mécanique statistique, prévoir le phénomène le plus probable. C'est ainsi que nous pouvons prévoir, pour prendre un exemple donné par Jeans, que de l'eau mise sur le feu se mettra à bouillir et ne se convertira pas en glace. Il n'est

1. On pourrait objecter que l'univers est peut-être fini ; nous n'en savons rien : mais, si l'espace est infini, comment éviter la dissémination indéfinie de la matière ? Les philosophes qui se refusent à croire à l'infini actuel trouveront sans doute d'ailleurs l'argument aussi fort, lorsqu'on se borne à constater le nombre extraordinairement grand des équations.

cependant pas *impossible*, au sens absolu du terme, mais seulement hautement improbable que l'eau mise sur le feu se gèle, de même que la reconstitution de la Bibliothèque nationale par une armée de singes dactylographes ne peut pas être déclarée impossible, mais seulement hautement improbable. Le « miracle de l'eau glacée sous l'action du feu » est exactement comparable au « miracle des singes dactylographes » ; il est probable et non certain, que ces « miracles » ne se produiront pas. Sans doute, il n'est pas douteux que ce degré de probabilité doit être, pour tout homme sensé, confondu avec la certitude ; cette certitude n'a pas cependant la valeur absolue qu'un philosophe déterministe doit attribuer à son déterminisme : car, si petite que soit la part faite à la liberté dans le monde, il y a un abîme entre un monde où cette part existe et un monde d'où elle est exclue.

Les remarques que nous avons faites sur la multiplication extraordinairement rapide par les chocs des déviations extrêmement faibles des molécules sont un des exemples du fait de l'importance des modifications qui seraient produites dans l'avenir de l'univers par un accroc, si faible et si localisé qu'il soit, au déterminisme universel.

Si, dans l'exemple de Jeans, l'eau mise sur le feu se congelait, ce ne serait pas bien entendu, pour le physicien atomiste, un phénomène réellement miraculeux, c'est-à-dire contraire aux lois scientifiques, mais seulement une coïncidence surprenante dont un hasard véritablement exceptionnel l'aurait rendu témoin ; le déterminisme atomistique ne subirait pas la plus légère atteinte, puisque ce que nous savons des mouvements atomi-

ques rend ce phénomène possible, bien que très hautement improbable. Ce qui serait atteint c'est ce que l'on peut appeler le déterminisme à notre échelle, c'est-à-dire le déterminisme qui nous est nécessaire pour que le monde nous apparaisse comme intelligible.

C'est, au fond, à cause de ce déterminisme-là que notre raison est naturellement portée à croire au déterminisme caché, mais absolu, des phénomènes moléculaires ; il est assez remarquable qu'il puisse y avoir contradiction entre les deux déterminismes.

113. — Cette contradiction étant constatée, il paraîtra peut-être moins étrange d'imaginer qu'il puisse y avoir, çà et là, dans le monde, des commencements absolus sans que ces commencements absolus soient incompatibles avec le déterminisme scientifique. Je dois avouer que cette conception de commencements absolus m'inspire une grande répugnance, comme à tous ceux, je pense, qui ont quelque éducation scientifique ou même simplement rationnelle; mais ma répugnance à l'égard de la théorie de la conscience épiphénomène[1] n'est pas moins grande et cette antinomie est le seul problème philosophique qui n'ait jamais cessé de me préoccuper. Je ne prétends pas l'avoir résolu ; mais l'antinomie me paraît tout de même moins embarrassante depuis que j'ai réfléchi sur le rôle du hasard dans les lois scientifiques ; c'est pourquoi j'ai essayé de résumer ici ces réflexions, si imparfaites qu'elles soient.

1. Voir Jules Tannery : *Revue du Mois* du 10 août 1906 (t. II, p. 129) et *Science et philosophie*, ch. III.

L'un des faits qui me paraissent le plus difficilement explicables dans la théorie de la conscience épiphénomène, c'est que le seul jeu du mécanisme matérialiste ait pu produire des livres dans lesquels sont exposées des idées abstraites et des sentiments complexes, livres dont l'action sur la nature peut être si considérable, par l'intermédiaire de l'homme. C'est là un « miracle » assez analogue au « miracle des singes dactylographes », mais plus invraisemblable encore ; l'existence de ce fait et sa persistance à travers les âges me paraissent incompréhensibles, si l'on n'admet pas que la raison abstraite de l'homme peut avoir quelque influence sur la marche de ses pensées et par suite sur ses actes. Les mouvements du cerveau qui correspondent à la pensée humaine sont certainement d'une complication comparable aux mouvements des molécules d'un gaz ; les uns comme les autres de ces mouvements sont scientifiquement déterminés lorsqu'on les observe à notre échelle, c'est-à-dire suivant des lois que nous ne connaissons pas toutes, mais que nous pouvons espérer connaître de mieux en mieux chaque jour. Mais ce que nous ne pourrons jamais analyser complètement, c'est le détail de ces phénomènes à l'échelle moléculaire ; il nous sera donc toujours impossible de savoir dans quelle mesure le déterminisme de ce détail est rigoureux. Si, d'ailleurs, au bout de millions de générations, l'homme, infiniment plus intelligent que nous, arrivait par des procédés que nous ne pouvons pas concevoir, à étudier par le dehors tous les phénomènes qui se produisent dans notre cerveau, ce progrès n'aurait pu se réaliser qu'après que le

cerveau de ce surhomme serait devenu infiniment plus complexe que le nôtre et par suite aussi difficile à connaître pour lui que le nôtre l'est pour nous.

114. — Il ne nous est pas possible de concevoir par quel moyen notre raison pourrait agir sur le mécanisme intime de notre pensée. On ne peut faire là-dessus que des hypothèses et des conjectures ; nous ne savons même pas comment notre conscience nous fait connaître, sinon ce mécanisme, du moins *quelque chose* d'équivalent à ce mécanisme. L'étude des phénomènes insconscients ou subconscients, des singularités si nombreuses de la psychologie pathologique, du rôle de l'intuition dans la vie pratique et dans la recherche scientifique[1] conduiraient à admettre que le mécanisme de notre pensée est infiniment plus complexe que la conscience claire que nous avons habituellement de cette pensée ; parfois, notre conscience arrive à saisir et à utiliser des portions de ce mécanisme qui lui sont habituellement cachées. C'est ainsi, pour prendre un exemple aussi simple que possible, que le calcul détaillé des combinaisons possibles et même le souvenir net de tous les coups antérieurement observés permettraient à un joueur de whist ou d'écarté de jouer d'une manière conforme à la probabilité maximum de gain. A défaut de l'examen détaillé de ces calculs ou de ces souvenirs, impossible en un temps très court, le joueur habile utilise son intuition, qui n'est peut-être

1. Voir E. Borel, L'intuition et la logique en mathématiques (*Revue de métaphysique et de morale*, 1907, p. 273) et L'évolution de l'intelligence géométrique (*ibid.*, 1907, p. 747).

que la conscience un peu vague d'un état de son cerveau trop complexe pour être perçu dans les détails. Il en est sans doute de même dans des cas plus compliqués et c'est en ce sens qu'il est légitime d'opposer, comme on l'a fait souvent, l'esprit de finesse à l'esprit géométrique. Ces remarques subsisteraient également d'ailleurs, il est à peine besoin de l'observer, dans l'hypothèse de la conscience épiphénomène. Ce qui me paraît beaucoup moins compréhensible dans cette hypothèse, c'est l'accord du monde avec notre raison ou si l'on veut, le fait que la connaissance de la vérité est avantageuse à l'évolution humaine.

Il ne s'agit pas de l'idée mystique de l'excellence de la vérité, mais du fait que l'évolution de l'espèce humaine et sa prédominance dans la lutte pour la vie ont été singulièrement aidées par l'acquisition des vérités vulgaires et des vérités scientifiques. Je sais bien que certaines personnes pensent[1] que la connaissance de la vérité suprême, c'est-à-dire la croyance au déterminisme absolu et à la croyance épiphénomène, aura pour conséquence de rendre impossible la vie de l'humanité. S'il en est ainsi, c'est une raison pour douter de cette vérité suprême, car ce serait le premier exemple de l'influence néfaste de la vérité sur l'évolution humaine. Ceux qui croient à un déterminisme rigoureusement matérialiste seraient d'ailleurs logiques avec eux-mêmes en n'éprouvant aucune répugnance à l'égard des mensonges dont la valeur sociale leur apparaît incontestable ; du moment que ces mensonges

1. Voir les ouvrages de M. Le Dantec et notamment son livre sur l'*Athéisme*, dont les conclusions pourraient presque se résumer par la parole célèbre : il faut une religion pour le peuple.

sont momentanément utiles, on doit les respecter et les répandre.

Ceux, au contraire, qui pensent que l'évolution de l'univers ne se poursuit pas suivant un plan immuable, mais que cette évolution peut être en quelque mesure influencée par la raison humaine, ceux-là doivent souhaiter que l'accord soit chaque jour plus complet entre cette raison et le monde qu'elle aspire à diriger ; ils doivent donc s'efforcer d'éclairer la raison de leurs semblables, avec la conviction que la vérité est toujours préférable à l'erreur et au mensonge et ne peut jamais être mauvaise. Cet idéalisme optimiste doit être la doctrine de ceux qui croient à la raison humaine, au progrès et à l'excellence de la vérité, le matérialiste pessimiste devant être laissé à ceux dont le pragmatisme s'accommode de toutes les erreurs traditionnelles, du moment qu'ils croient leur apercevoir des avantages immédiats.

115. — Dans le domaine scientifique même, l'introduction des méthodes statistiques a conduit à modifier la conception traditionnelle du progrès. Réaliser un progrès, ce fut d'abord expliquer le compliqué par le simple, ramener à la loi de Newton, à l'attraction en raison directe des masses et en raison inverse du carré des distances, la complication des mouvements des astres. Cette notion du progrès subsiste, mais doit être complétée ; la loi de Newton nous paraît toujours comme une étape nécessaire et indestructible de la pensée scientifique, mais elle ne nous satisfait plus entièrement par sa simplicité et nous admirerions à l'égal d'un Newton le physicien qui *expliquerait* cette loi par une théorie

statistique d'où elle découlerait comme une conséquence infiniment probable de l'irrégularité de certains mouvements extrêmement complexes.

116. — L'une des conséquences des théories physiques modernes sur lesquelles on a le plus discuté est ce que l'on a appelé la doctrine de la dégradation de l'énergie.

Le second principe de la thermodynamique, ou principe de Carnot, a été énoncé par Clausius sous la forme précise que l'entropie d'un système isolé ne peut pas décroître : par une extension hardie à l'univers entier, on en a déduit souvent cette conséquence d'une portée bien plus générale : l'entropie de l'univers tend vers un maximum, d'où l'on peut conclure aussi que l'évolution est irréversible. Si l'on est pessimiste, cette irréversibilité de l'évolution peut signifier que l'univers tend vers un état de mort thermique, dans lequel l'énergie *utilisable* aura disparu ; si l'on est optimiste, on peut se contenter de cette conséquence en un certain sens consolante, que le progrès n'est pas un vain mot, puisque les états antérieurs du monde, une fois dépassés par l'évolution, ne se retrouveront plus jamais. Ces diverses conséquences dépassent beaucoup, il n'est pas besoin de le dire, la portée du second principe ; en particulier, il faudrait d'abord résoudre la question de savoir si les réserves d'énergie utilisable sont finies, dans un système fini, et *a fortiori* dans le système infini qu'est peut-être l'univers. A plus forte raison dépasse-t-on largement les résultats effectivement acquis à la science si l'on donne au second principe l'énoncé le plus général qu'on obtient en interprétant tout d'abord les

résultats les plus simples dans le langage de la mécanique statistique et en étendant ensuite l'énoncé ainsi obtenu aux cas les plus généraux. On arrive ainsi à la loi générale de l'évolution, d'après laquelle l'univers va constamment des états les moins probables vers les états les plus probables. Il est évidemment impossible de démontrer l'exactitude ou l'inexactitude de cet énoncé, tant que l'on n'a pas défini d'une manière précise la probabilité. On peut simplement observer que l'énoncé est exact dans les cas simples où la probabilité a pu être définie, et induire de là qu'il doit être possible de définir dans tous les cas la probabilité de telle manière que l'énoncé subsiste.

Je voudrais dire ici quelques mots des remarques de Boltzmann sur l'application du deuxième principe à l'univers. Comme le dit fort justement Boltzmann, « assurément personne ne prendra de telles spéculations pour d'importantes découvertes, ni pour le but le plus élevé de la science, comme le faisaient les anciens philosophes. Mais il n'est pas certain qu'il soit juste de les tourner en dérision et de les regarder comme tout à fait oiseuses ». Boltzmann développe une conception mécanique de l'univers, dans laquelle il se produit, çà et là, des passages d'un état plus probable à un état moins probable, de sorte que, pour l'univers entier, l'irréversibilité n'existe pas. Cette conception est rigoureuse au point de vue abstrait si l'univers est un système mécanique pouvant être défini par un nombre fini de paramètres dont le champ total de variation est fini. Admettons, pour un instant, que nous puissions accepter cette image pour l'univers que nous voyons, c'est-à-dire que nous puissions fixer un

nombre très grand R, tel qu'il n'y ait jamais rien à l'extérieur de la sphère S de rayon R ; cette sphère S sera notre univers ; l'évolution de cet univers sera, d'après un théorème de Poincaré, aussi voisine que l'on veut d'une évolution périodique et, dans des périodes immensément longues, les phénomènes en contradiction avec le second principe y seront aussi fréquents que les phénomènes en accord avec ce principe. En laissant même de côté les difficultés — cependant à mon avis insurmontables — entraînés par l'hypothèse que rien ne sort de la sphère S, il faut observer que la conclusion n'est rigoureuse qu'autant que nous supposons *absolue* l'inexistence de toute action extérieure à S. Imaginons, avec O. Chwolson[1], une sphère S_2 dont les dimensions par rapport à S seraient celles de S par rapport à un atome, puis une sphère S_3 qui serait à S_2 ce que S_2 est à S, et ainsi de suite jusqu'à une sphère S_n dont l'indice n serait égal à un million. Pour que l'application à S de la théorie mécanique de la quasi-périodicité due à Poincaré fût légitime, il faudrait que nous fussions assurés qu'il n'y a pas, aux confins de S_n, quelque univers S' de mêmes dimensions que S bien que probablement très différent de S et pouvant, dans le cours des temps, agir sur S. Car la durée des temps nécessaires pour l'application du théorème de Poincaré est tellement longue qu'une rencontre de S avec S' serait infiniment probable, bien avant que ces temps fussent écoulés. Ceci revient à dire qu'il est au moins aussi vraisemblable de supposer que les lois de notre univers seront complètement

1. *Scientia*, t. VIII, 1910, pages 41 (texte) et 45 (suppl.).

modifiées par une combinaison avec un autre univers (actuellement infiniment plus éloigné de lui qu'un atome situé sur la Terre n'est éloigné d'un atome situé sur Sirius) que de supposer un changement de sens appréciable dans la variation de l'entropie.

En d'autres termes, l'évolution régulière vers des états de plus en plus probables me paraît, contrairement à la conception de Boltzmann, devoir être admise pour l'univers entier, du moment qu'on ne le regarde pas comme un système fini isolé pour toujours dans une portion finie de l'espace, de laquelle rien ne peut sortir, ni matière, ni énergie, ni rayonnement et dans laquelle rien ne peut entrer [1].

117. — La question se pose alors de savoir si l'on peut concevoir le commencement et la fin de cette évolution ; une telle question est difficilement soluble d'une manière entièrement satisfaisante, car notre esprit ne conçoit pas plus aisément l'infini du temps que l'infini de l'espace ; tout ce que l'on peut demander à une théorie cosmogonique, c'est de ne pas augmenter encore, par ses particularités, les difficultés inhérentes à toute question dans laquelle l'infini est en jeu.

C'est ainsi que certains physiciens et notamment Arrhénius, ont cherché à imaginer com-

1. On pourrait chercher à diminuer l'étrangeté (pour ne pas dire plus) de cette conception de l'univers fini en supposant que l'univers est périodique, c'est-à-dire se compose d'une infinité de parallélépipèdes rigoureusement identiques et par suite indiscernables, de sorte que tout se passerait comme si l'univers était fini et avait, avec une dimension de plus, une structure géométrique analogue à un rectangle dont on aurait collé deux à deux les côtés opposés, ce qui donne un tore. Mais tout cela est bien peu vraisemblable.

ment les conditions particulières dans lesquelles se trouvent les nébuleuses peuvent rendre possible la reconstitution des corps radioactifs, ainsi que d'autres infractions au second principe [1].

Je dois avouer que je suis assez sceptique en ce qui touche la possibilité de ces infractions ; l'évolution vers les états les plus probables me paraît être une loi générale, à laquelle il n'est pas possible d'échapper.

Maxwell a bien cherché à montrer comment un être très intelligent et très ingénieux pourrait, sans dépenser aucun travail, en ouvrant ou refermant simplement un orifice par une porte maniable sans effort, arriver à faire évoluer une masse gazeuse vers les états les moins probables : il suffirait que ce « démon de Maxwell » laisse passer dans un sens les molécules les plus rapides (c'est-à-dire les plus chaudes) et dans l'autre sens les molécules les plus lentes (c'est-à-dire les plus froides). De cette manière les deux masses gazeuses séparées par l'orifice, acquerraient spontanément une différence de température. Poussant plus loin la même idée, M. Jean Perrin s'est demandé si l'on ne pourrait pas concevoir l'utilisation du mouvement brownien, conséquence visible et accessible pour nous de l'agitation moléculaire ; il semble bien qu'il nous serait possible à la rigueur, au prix de beaucoup d'efforts, d'obtenir une très minime infraction au principe de Carnot. On peut se demander toutefois, si les divers mécanismes que l'on peut concevoir dans ce but n'exigeraient

1. On arrive ainsi à donner une base quasi scientifique à la fameuse doctrine du « retour éternel ».

pas forcément l'intervention constante d'une intelligence humaine, intervention qui suppose un fonctionnement vital, c'est-à-dire un accroissement d'entropie peut-être forcément plus élevé que la diminution obtenue par ailleurs.

Même en laissant de côté ces objections, l'argument du démon de Maxwell est simplement la constatation du fait que la raison humaine sait créer l'ordre là où le jeu spontané des forces mécaniques tendrait à augmenter le désordre. Tous les produits de l'industrie humaine, qu'il s'agisse de kilomètres de rails ou de bibliothèques remplies de livres dans lesquels le désordre apparent des caractères est en réalité un ordre d'essence supérieure, tous ces produits sont, *a priori*, infiniment peu probables. Mais la place qu'ils occupent dans l'univers est si peu de chose et d'autre part la dépense d'énergie utilisable nécessaire pour obtenir ces résultats, le désordre créé pour obtenir cet ordre, le charbon dissipé pour agglomérer ce fer, etc., sont, en fait[1], relativement si considérables, que l'on peut négliger cette création d'ordre[2], lorsque l'on envisage le problème cosmogonique ; on peut même négliger, semble-t-il, tous les phénomènes biologiques par rapport à l'ensemble de tous les autres phénomènes ; l'évolution de l'univers vers les états les plus probables, devient alors une loi absolument générale.

1. En toute rigueur, il n'est pas démontré que le perfectionnement des techniques ne permettrait pas d'obtenir, grâce à la raison, plus d'ordre qu'on ne crée de désordre ; c'est une question que je ne peux aborder ici.

2. On pourrait, à propos de cette création d'ordre par la raison, revenir sur ce que nous avons dit plus haut (n° 114) ; mais je dois me borner.

118. — Peut-on, en admettant cette loi générale de l'évolution, concevoir cependant la durée illimitée de cette évolution dans le passé et dans l'avenir, sans qu'à aucun instant il puisse être question de mort de l'univers? Ce n'est pas impossible : il suffit d'imaginer que le désordre, c'est-à-dire les différences de température, l'hétérogénéité de la matière, etc., s'accroît constamment à une échelle de plus en plus petite, tandis que l'ordre va sans doute en diminuant à l'échelle où le désordre s'accroît mais peut s'accroître et même se perfectionner à une échelle encore plus faible. En d'autres termes, la structure de l'univers devient de plus en plus fine; la faible fraction de l'énergie solaire qu'absorbe notre globe, y a créé les mines de charbon, très petites à l'échelle du soleil, mais qui représentent à notre échelle une provision considérable de matériaux ordonnés; la dissipation du charbon nous permet de créer de l'ordre à une plus faible échelle; il est vraisemblable qu'il se passe des phénomènes analogues aux échelles trop grandes ou trop petites pour nous être accessibles.

L'évolution de l'univers pourrait ainsi être conçue comme tendant à produire un état de plus en plus compliqué ne pouvant être perçu et utilisé que par des êtres de plus en plus petits. Comme il n'y a pas d'étalon absolu de longueur[1], un tel rapetissement n'a rien qui doive nous effrayer; il nous semble actuellement que des

1. On a proposé comme tel étalon les longueurs d'ondes lumineuses; ces longueurs d'onde sont relatives à nos yeux et à la nature des corps actuellement regardés comme simples. Le principe de relativité introduit un étalon absolu de *vitesse*, mais non de longueur ni de temps.

êtres de la taille des molécules, et surtout des êtres qui seraient par rapport aux molécules ce que nous sommes par rapport au soleil, sont bien peu dignes d'intérêt pour nous ; il n'y a aucune raison pour que la complication croissante de l'univers ne rende pas possibles, s'ils ne le sont déjà, de tels êtres avec une structure relativement plus complexe que la nôtre.

L'évolution de l'univers pourrait être ainsi regardée, au bout d'intervalles de temps immensément longs, comme une réduction en quelque sorte homothétique accompagnée peut-être d'une tendance de plus en plus grande à l'ordre dans le sens élevé de complexité réglée par la raison. Il semble tout au moins permis d'affirmer qu une telle conception ne contredit pas les données nouvelles que nous devons à la science du hasard.



TABLE DES MATIÈRES

PREMIÈRE PARTIE

LA DÉCOUVERTE DES LOIS DU HASARD

DEUXIÈME PARTIE

L'APPLICATION DES LOIS DU HASARD

TROISIÈME PARTIE

LA VALEUR DES LOIS DU HASARD

ÉVREUX, IMPR. CH. HÉRISSEY, PAUL HÉRISSEY, SUCC^{r}

551-20. — Coulommiers. Imp. Paul BRODARD. — 11-20.



